NEURAL CIRCUIT MECHANISMS OF STIMULUS SELECTION UNDERLYING SPATIAL ATTENTION

by

Nagaraj R. Mahajan

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

February 2020

Abstract

Humans and animals routinely encounter competing pieces of information in their environments, and must continually select the most salient in order to survive and behave adaptively. Here, using computational modeling, extracellular neural recordings, and focal, reversible silencing of neurons in the midbrain of barn owls, we uncovered how two essential computations underlying competitive selection are implemented in the brain: a) the ability to select the most salient stimulus among all pairs of stimulus locations, and b) the ability to signal the most salient stimulus categorically.

We first discovered that a key inhibitory nucleus in the midbrain attention network, called isthmi pars magnocellularis (Imc), encodes visual space with receptive fields that have multiple excitatory hotspots ("lobes"). Such (previously unknown) multilobed encoding of visual space is necessitated for selection at all location-pairs in the face of scarcity of Imc neurons. Although distributed seemingly randomly, the RF lobe-locations are optimized across the high-firing Imc neurons, allowing them to combinatorially solve selection across space. This combinatorially optimized inhibition strategy minimizes metabolic and wiring costs.

Next, we discovered that a 'donut-like' inhibitory mechanism in which each competing option suppresses all options except itself is highly effective at generating categorical responses. It surpasses motifs of feedback inhibition, recurrent excitation, and divisive normalization used commonly in decision-making models. We demonstrated experimentally not only that this mechanism operates in the midbrain spatial selection network in barn owls, but also that it is required for categorical signaling by it. Moreover, the pattern of inhibition in the midbrain forms an exquisitely structured 'multi-holed' donut consistent with this network's combinatorial inhibitory function (computation 1).

Our work demonstrates that the vertebrate midbrain uses seemingly carefully optimized structural and functional strategies to solve challenging computational problems underlying stimulus selection and spatial attention at all location pairs. The neural motifs discovered here represent circuit-based solutions that are generalizable to other brain areas, other forms of behavior (such as decision-making, action selection) as well as for the design of artificial systems (such as robotics, self-driving cars) that rely on the selection of one among many options.

Thesis committee

Advisor and First Reader: Dr. Shreesh P. Mysore

Second Reader: Dr. Hynek Hermansky

Committee members: Dr. Mounya Elhilali, Dr. Ralph Etienne-Cummings

Acknowledgments

My PhD thesis would not have been possible if not for the support, both professionally and personally from many people in my life.

I am very grateful to my advisor Dr. Shreesh P. Mysore for his immense support and mentorship through the course of my PhD, and for providing me a home in his lab. I had the fortune of meeting Shreesh through Dr. Hynek in my first year. He was kind enough to offer me a rotation in his lab despite my lack of experience in neuroscience. I had a great time during the rotation as he was always available for discussions, and was willing to spend hours together on meetings to answer even basic neuroscience questions. This cemented my decision to join his lab over the other labs I was considering at that point.

Over the years, the excitement and passion with which he spoke about research has been extremely contagious. I have learnt so much through our uncountable meetings and discussions that many times lasted for several hours. Through these discussions, he not only shared his knowledge about the field of neuroscience, but also imparted his wisdom about his experiences going through various phases of an academic and research career.

I am thankful for the training he provided me with for performing experiments and surgeries. Rather than relying on some preexisting data to model with, this gave me the freedom to design and perform experiments to answer questions that were of direct relevance to our dynamically evolving research findings. For this reason, I am glad that I pursued a combination of experimental and computational neuroscience for my PhD. As we worked through figuring out the computational principles associated with some of our exciting (at least to us) experimental findings, I learnt from him how to frame a fundamentally important research question and answer it methodically. Working on both the 'combinatorial inhibition' and 'categorization' projects has been such an exciting, fulfilling and enjoyable experience for me. From Shreesh, I have also learnt the value of writing and presenting results in a compelling manner. Additionally, he gave me extremely valuable feedback on numerous research presentations across the years, and my ability to present a talk has significantly improved over the years, thanks to his feedback and time.

When I joined Shreesh's lab, I was quite surprised at his willingness to accept me, someone with no experimental background, into his lab and to train me in doing some really challenging neuroscience experiments. Looking back after all these years though, I realize that it was just a mere reflection of who Shreesh is as a person, and of his scientific and mentorship principles. As I ventured into a totally new field, there were obviously numerous times where I doubted myself. But his endless encouragement and patience helped me instill a strong belief in my decision to explore the magical world of neuroscience, something I have come to love so dearly. I will be eternally grateful that he was willing to take a chance on me and invest in me, for I have benefitted and learnt a lot from it. As I culminate my thesis, I do feel that I have not only found a mentor and an advisor, but also a friend in Shreesh to whom I can reach out to, as I continue pursuing research in the coming years of my life.

I also have a lot to thank to Hynek for. When I told him in my first year that I was interested in neuroscience, he was extremely supportive and encouraged me to follow what I was truly interested in, even though I had come to Hopkins with an intention of working with him for my PhD. He not only gave me a neuroscience project to explore the area to see if it matched my interests, but also met numerous neuroscience professors along with me to see if we could work on potential collaborative projects. I am grateful for his mentorship and advice across the years related not just to science, but also career. I enjoyed working with him on projects related to STRF analyses. I thank him for his guidance and advice, both in that project and my thesis work.

I thank Dr. Mounya Elhilali and Dr. Ralph Etienne-Cummings for serving on my thesis committee; Dr. James Knierim, Dr. Noah Cowan and Dr. Cynthia Moss for serving on my gradual board oral committee; and Dr. Sanjeev Khudanpur and Dr. Trac Tran for serving on my department qualifying exam committee. Their comments and feedback at various points of my PhD were extremely valuable in putting together this thesis. I reached out to Dr. Rene Vidal and Dr. Amitabh Basu to discuss some aspects of my research, and getting their perspective was very helpful in thinking about the broader impact of my project. I also thank Dr. Hita Adwanikar for giving feedback about my projects during lab research presentations. I thank Dr. James Knierim for his valuable and patient feedback on our papers.

My interest in pursuing a PhD and research in general is credited to the wonderful set of professors I had the fortune of working with during my bachelor's degree at IIT Hyderabad. My undergraduate thesis advisor, Dr. Sri Rama Murty's ability to dissect a complicated problem and solve it with a systematic approach is something I will only hope to achieve one day. Dr. Soumya Jana encouraged me to not be overwhelmed by difficult problems. As I struggled through the open-ended research problem he gave us as course projects, the meetings and conversations with him taught me to break down big problems into smaller solvable ones, and to celebrate any progress as a victory. My interest in neuroscience and the biomedical field broadly began due to my work with Dr. Sumohana Channappayya on using image processing techniques for analyzing retinal scan images. That project made me realize that I could use my background in DSP in interdisciplinary fields, and evidently, that has had a lasting impact on me as I culminate this thesis in neuroscience. The experience I gained by working with the three of them during my formative years will stay with me for the rest of my life. I also fondly remember the courses by Dr. Balasubramaniam Jayaram. I hope that if I end up becoming a teacher one day, I could teach like him.

Staying far away from family in a different country would have been infinitely harder if not for the support system from lab mates and friends at Hopkins. Over the years, I have learnt so much from my brilliant, extremely smart and kind lab mates: Wen-Kai, Margo, Juliana, Hannah, Ninad, Garima and Barbara. It has been absolutely great to share my PhD journey with you guys and learn from your journeys and projects. I have had numerous discussions with all of them about my research work, especially with Ninad, Wen-kai and Margo. Coffee and lunch conversations became a regular occurrence once Ninad joined the lab. I have discussed my research with him quite extensively, and getting his thoughts on it always made me think harder about my work. During the initial years, as the lab was being setup, figuring out and learning from several rig, analysis and experimental related issues along with Shreesh and Margo was a great scientific learning experience for me. It not only taught me a lot about electrophysiology, but also gave me a clear understanding of the TDT system which came in handy to find and implement solutions when many issues came up during the later parts of the PhD.

I am also very thankful to David who maintained the owl runs and fed the birds, and Jim who made any tools that was required for the owl experiments. With their assistance, I could focus better on my experimental data collection. I thank the admin staff from both the ECE and PBS departments: Julie, Debbie, Rebecca, Belinda, and Ruth for their help with administrative and non-administrative matters.

At a crucial juncture during my PhD, I attended the fantastic Neural Systems & Behavior course at MBL Woods Hole. There, I had the fortune of learning from all the brilliant and kind instructors (too many to list here). I got to work on numerous model organisms, learn experimental and analysis techniques, and work on projects that I would never otherwise have worked on. This course helped me frame my research questions with the broader context of the field in mind. Apart from everything I learned there, I also made extremely good friends, some of whom I stayed in constant touch with, like Poonam and Yerko. I have reached out to them to talk about my PhD joys and sorrows, and hear about theirs as we navigated grad school life together in different parts of the world. I still cherish what I learnt in those two months and the friendships I gained there. I am extremely thankful to Shreesh who suggested that course and encouraged me to apply for it.

Outside of lab, I was kept sane thanks to a wonderful group of friends over the years: Vivek, Shourya, Varun, Mithra, M, Akshay, Garima, Ranjani, Purnima, Harini, Ramanujan, Vishwa, Benita, Shruthi, Vimal and Ramanujan. A common stress buster was hanging out with them, playing board games, badminton, and tennis. During the first few years, I enjoyed volunteering for AID JHU and working alongside Nishikanth, Mustafa and Santosh.

During undergrad, I had a great set of friends with whom I formed study groups and worked on projects with (Ravi Chandra, Aditya, Aniket, Maaz, Shabaz and Anil). It would not be too farfetched to say that our shared interest in research and solving challenging problems directly contributed to me applying for and pursuing PhD. I also thank folks from LIGAF and friends from MES for their friendship across the years.

Coming from a joint family, I have a lot of people to thank. Being born in a village in rural India with poor educational facilities, my journey to pursing a PhD is one that is unfortunately not common, and one that should not have been possible. Anything I am today is only because of my uncle and aunt, Sanjeev and Pushpa Mahajan, and my parents, Rajappa and Sridevi Mahajan. When I was 4 years old, my uncle and aunt decided to bring me to Bangalore along with them after their marriage, and raise me as their own child. They provided me with all the love, the best of education, and nurtured my growth and have been no less than parents to me. My parents have been extremely loving, supportive, patient and have always encouraged me to pursue things I was passionate about throughout my life. The four of you have encouraged me to pursue whatever I wish, inspired and motivated me to not be fazed by difficult things in life, and stood as a rock behind me in difficult times. Your love and care towards me has been immeasurable and I count my blessings every day for it. Your hard work, dedication and sacrifice are something I cannot even begin to fathom. Often people search for inspiration in their lives. I have been fortunate in that I do not have look too far with the four of you in front of me. I am also indebted to my grandmother, one of the strongest persons I have known in my entire life, and my granduncle for the emphasis he placed on education even in difficult times in our family. The fact that I am able to pursue this opportunity today is only because of the sacrifices and hard work across the two generations in my family is not lost on me. I cannot even begin to thank all of you, but I would like to dedicate this thesis to the six of you as a small token for all the love you have given me and everything you have done for me.

Regular chats with my loving siblings Apoorva, Ananya, Ambica (and my brother in law, Sunil) and Ravi (and my sister in law, Asha) have always kept me grounded, and helped me not to take anything too seriously in life. Any stress or worry at the end of the day just immediately vanishes when I speak to my adorable nephews, Neel and Aarav. Additionally, I would also like to thank my cousins, uncles and aunts for their support over the years. I am extremely lucky to have a loving, supportive and a brilliant partner in Pravallika. She has been a constant in my life through thick and thin, day and night. It is difficult to imagine what my life would be without you in it. Many of the things (too many to elaborate here) during my PhD would not have been possible without your constant love and support. Apart from being an emotional pillar during some extremely difficult phases in my personal life, she keeps inspiring me through her patience and hard work in my professional life. Her kindness, empathy and patience in every aspect of life inspires me to strive and learn from her every day. Life just seems so right with you in it. Thank you for everything you have brought into my life!

All the wonderful people above have played such a vital role in making this journey possible. However, I am acutely aware that I was able pursue this PhD only because I was fortunate enough to get the opportunities that they gave me in every phase of my life. I would sincerely like to dedicate this thesis to folks like me from my village Partapur, and other villages across India (and other parts of the world), who probably would have done a better job than I did, but just weren't fortunate enough to have gotten the opportunities that were given to me in my life.

Table of Contents

List of Figures

1. Introduction

1.1 Selective attention: Background and early theories

At any given instant, our brain is bombarded by a vast amount of information. It is estimated that the human retina by itself sends an astonishing 10^7 to 10^8 bits per second to the brain [1]. Not all of this information is relevant to guide our behavior. More importantly, our brain has limited energetic and computational resources making it impossible to analyze all of this information [2-6]. Much of the information received through our sensory organs is 'ignored', and only relevant pieces of information are selected and processed further. Attention is the process of selecting the most relevant information at each instant, and preferentially processing it to guide our behavior. A common analogy used to depict this concept is that attention is an informational bottleneck (Figure 1.1), allowing only a select subset of all the information received to enter our short-term working memory, and to guide our behavior.

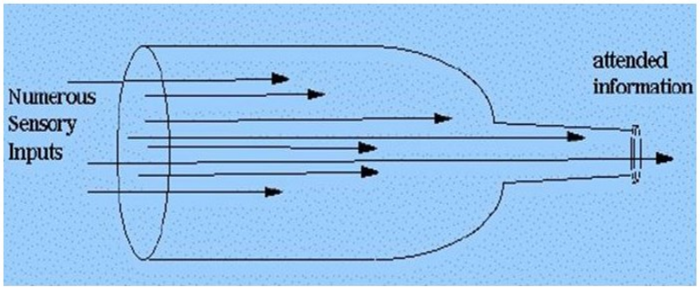


Figure 1.1: Attention as an information bottleneck. The brain receives multiple pieces of information through sensory inputs, but only a subset of it is selected and passed through the bottleneck for further processing. Figure adapted from https://www.simplypsychology.org/attention-models.html

Several pieces of experimental evidence support the bottleneck theory of attention. Some of the earliest support comes from experiments by Donald Broadbent, an influential psychologist, who proposed the 'filtering theory' of attention in 1958 based on a dichotic listening task [7]. He had participants wear headphones and listen to two separate messages in each ear. They were instructed to pay attention to the message in one ear and ignore the message in the other. After the task, he probed the participants about the two messages and found that subjects performed poorly when asked for details about the message in the ignored ear, but did very well on details regarding the message in the attended ear. Based on this, he proposed that whereas information relevant to behavior passes through for further processing, that which is irrelevant is filtered out, ignored and 'lost' (Figure 1.2A).

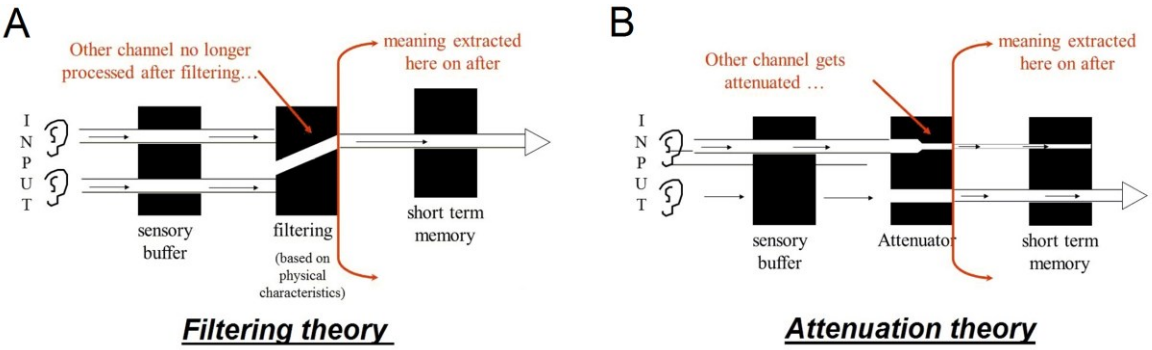


Figure 1.2: Theories of attention.
(A) Filtering theory of attention: The selected piece of information is passed through the bottleneck and other irrelevant pieces are blocked out. (B) Attenuation theory of attention: The selected piece of information is passed through the bottleneck and the other irrelevant pieces are also passed through, but only after initially attenuating them. Figure obtained from **https://www.simplypsychology.org/attention-models.html**

However, Broadbent's filtering theory fell short in the aspect of what happens to the ignored pieces of information. In contrast to his original proposal, follow up experiments revealed that participants were capable of recollecting information from the unattended ear too depending on what the message in that ear was [8-10]. For instance, if the message was the name of the participant, or if the message was relevant to the message in the attended ear, the participants were able to recollect that information. If unattended information is lost, how are participants able to recollect the information from the unattended ear?

To explain this, Anne Treisman proposed the 'attenuation theory' of attention. She suggested that the ignored pieces of information are also passed through the bottleneck, but only after their representations are 'attenuated' (Figure 1.2B; [11]). The irrelevant pieces of information are not fully lost in the process of filtering, but are available in case they may become relevant later. A standard example of this is the 'cocktail party problem'. Imagine you are at a cocktail party and talking to your friend in the midst of all the chatter in the hall. Humans do a very good job of ignoring the background chatter and specifically listening to the person they are speaking to. But imagine if someone were to call your name. Your brain is immediately able to pay attention to such a stimulus despite the ongoing background chatter. Such behavioral phenomena are supported by the attenuation theory of attention but not the filtering theory of attention. Through the selective processing capability that attention bestows, our brains are able to navigate through the cacophony of information at each instant and reliably guide behavior.

Attention can be divided into distinct types depending on what drives it and how it is deployed.

(a) Based on whether or not orienting movements of the eyes/ears and head are involved during attention, it can be classified as 'overt' vs. 'covert' attention [12]. Overt attention is the process of allocating attention to a particular location through orienting movements of sensory organs that bring the target to the center of the 'gaze'. Covert attention is the process of directing attention without explicit physical movements of the sensory organs, but by mentally shifting attention on the target.

(b) Based on whether an observer is allocating attention to a target voluntarily or reflexively, attention can be either ‘endogenous’ (also called top-down attention) or ‘exogenous’ (also called bottom-up attention) [12]. Endogenous or top down attention is when an observer willfully directs attention towards a target. Exogenous or bottom up attention is driven by the (sudden) appearance of a novel, relevant stimulus in the surroundings. Whereas the former is goal driven, the latter is driven by external stimuli in the sensory world.

(c) Based on whether attention is allocated in a spatially dependent manner or not, it can be divided into ‘spatial attention’ versus other forms such as ‘feature-based attention’, ‘object-based’ of ‘non-spatial’ attention [12]. Spatial attention is the direction of attention (either covert or overt) to a particular location in the world independent of the features of the target occurring there. Feature-based attention is the direction of attention to a particular feature (such as color, contrast etc.) independent of spatial locations. Object-based attention is the direction of attention to a particular object, such as a house when overlapping translucent images of houses and trees are presented. Non-spatial attention includes the direction of attention to mental concepts, for instance attending to the memory of yesterday’s movie while ignoring the ongoing lecture in class.

Even though attention is classified into multiple types based on these various factors, they are not mutually exclusive. For instance, studies have shown that overt and covert attention can interact: the allocation of covert attention is typically (though not always) followed by and aids overt attention [12-14]. Endogenous and exogenous control of attention have been shown to interact with and facilitate each other during vigilant behavior [15]. Similarly, feature based attention and spatial attention can interact with each other (for instance, searching for someone wearing a black shirt on one side of a crowded room) [16, 17].

In sum, it is easy to see how selective attention can be essential for survival and adaptive behavior. Humans and animals continuously navigate situations that are potentially dangerous and contain threats to their survival (imagine prey animals in the wild looking out for potential predators; or imagine crossing the street and looking out for oncoming vehicles). In the face of the brain's limited processing capacity, attention becomes an especially important and essential computation to ensure the survival of species [18]. Detailed investigations over the past half-century or so have revealed several insights into the neural correlates of attention and the ways in which selected information is preferentially represented and processed. I will summarize these next.

1.2 Effects of attention

Attention operates in multiple ways to improve the signal-to-noise ratio (SNR) of the representations of the selected stimulus. In the last few decades, numerous studies have highlighted the effects of attention on neural responses. It is well known that attention modulates activity by increasing the firing rates of neurons encoding for the attended stimulus (also called as an attentional gain of responses). Such gain modulation of responses results in a direct increase of SNR of the attended stimulus [12, 19-24].

Neurons have inherent variability in their responses to repeated presentations of the same stimulus. Another mechanism through which attention improves the SNR is by reducing such variability in a neuron's response across trials (quantified via the so-called fano-factor), when attention is directed to the preferred input (or receptive field, RF) of the neuron. Such reduction in response variability can help a downstream decoder to read out the activity more accurately [25-27].

Additionally, attention also reduces correlation in responses among neurons encoding the attended stimulus (this is termed as noise correlation). Neurons encoding for the same stimulus typically share some variability. Such reduction in correlated variability makes information theoretic sense, since it reduces the mutual information among neurons and makes their responses more independent from one another. This increases the overall informational capacity of the neuron population. Such decorrelation also aides a downstream decoder to read out the information from a population of neurons more accurately [28-31].

The above mechanisms of attention primarily operate on the spiking activity of neurons. Attention also modulates non-spiking neural responses, specifically, the local field potential (LFP) – the low frequency voltage changes representing the aggregate or population activity within a small patch of neural tissue. Attention has been shown to increase the synchrony in these low-frequency oscillatory responses of neuronal populations encoding an attended stimulus, particularly in the gamma band frequency (25-140 Hz) [32, 33]. Such synchronization can ensure that a downstream decoder is able to read out the activity related to the attended stimulus more efficiently, since it receives stronger input (simultaneous input from a larger number of neurons) [33]. This leads to gain modulation of the activity of downstream neurons, in-turn increasing the SNR of their responses [34]. Studies have also reported that along with an increase in the synchrony in the gamma band frequency, there is a reduction in the alpha and beta band synchrony between neurons [33, 35, 36]. Typically, synchrony in those bands is associated with inhibitory activity. Some studies have found an increase in the synchrony in alpha and beta bands among neurons encoding for distractors, thus suppressing them [37, 38].

Finally, attention increases the efficacy of communication between neurons across brain areas encoding for the attended stimulus. While paying attention, it is important that the brain area encoding the task rule (internal goals) efficiently communicates the rule to the parts of the brain that deploy attention. Studies have found that neurons with overlapping RFs in the PFC and visual areas such as V4 have higher phase locked spike-spike and spike-LFP activity. This increase in phase locking ensures that the activity from the first area reaches the second area during a time window that is most efficient at driving neurons in the second area, thus enhancing the communication between the two areas [39]. One key difference between the gamma band synchronization in the local areas vs. cross-area synchronization is the phase. In local areas, the phase lag is zero since all neurons spike together. In cross-area synchronization, the phase lag is non-zero reflecting the time required for spikes to travel form the first area to the second. Multiple studies across species [32, 36, 40-45] have provided evidence of increased synchrony as a correlate of increased interareal communication. Note that such local as well as cross-area increases in gamma band activity are not unique to attention; they also play a role in other contexts: respectively, encoding sensory stimulus properties [46], and encoding other behavioral states (for instance, memory [47, 48]).

It is clear that these neural mechanisms of selective attention can help highlight and improve the representations of the information that is most important/relevant to behavior at any instant. However, an essential question underlying the direction of attention is, “How does the brain decide and select at each instant, which piece of information is most relevant, and which pieces are unimportant and to be ignored (or attenuated)?” In other words, what are the neural computations that subserve selection for attention, and how are they implemented in neural circuits? Despite the long history of the neurobiological research on selective attention, this fundamental question relating to the control of the attentional locus remains largely unanswered.

1.3 Competitive selection, salience maps and priority maps

A framework for thinking about how competition among the various pieces of information and selection of the target of attention might occur comes from theoretical work on the salience (priority) map. This concept arose from the study, specifically, of selective spatial attention, i.e., the problem of identifying the location in the world that contains the most important information for behavior at any instant, but is now thought to apply more generally.

1.3.1 What is a saliency map?

The concept of a saliency map was first described in a seminal paper by Koch and Ullman [49]. Saliency was defined to be a property that depends on the relative distinctiveness (or conspicuousness) of 'targets' in a scene. Imagine a simple example where you are driving down the road and you approach a red traffic signal. The red traffic light would likely be the most conspicuous target (at that instant) that guides behavior - you bring your car to a stop. However, imagine instead that you hear the sound of tires screeching nearby just as you are slowing down. The sound of screeching tires would supersede the traffic light as the most conspicuous target and guide behavior - you move over to a parking lane to avoid a potential collision. So the same target, the traffic light in this example, is the most salient in one context and not in another, because of the relative distinctiveness of the objects in the surrounding. Salience in a scene is the property of assigning a value (a number) that represents the relative 'importance' (or conspicuity) of various targets in that scene. The core idea was that targets that are more conspicuous and stand out in comparison to others are selected and passed through the attentional bottleneck, i.e., selected. Consequently, the computation of the relative importance and subsequent selection of the targets are generally thought to be a 'pre-attentive stage' of information processing. The 'post-attentive stage' of processing involves enhancing the representations of the selected stimuli (using attentional mechanisms described in section 1.2) and using the enhanced representations to further guide behavior.

1.3.2 Implementation of the saliency map as a computational model

Koch and Ullman [49] proposed that the saliency map is obtained from a number of spatially aligned maps of features such as color, orientation, contrast, direction of movement etc. (Figure 1.3A). Each of these maps encodes space topographically (i.e. nearby 'units' or 'neurons' in the map encode for nearby locations in sensory space, and there is a systematic one-to-one ordered relationship between the locations in the sensory space and the units encoding for those locations in the feature maps). In each of these feature specific maps, locations in the maps compete with one another and the 'most important' location for that feature map is identified based on which location wins this competition. These individual feature dependent maps are then normalized and combined together to obtain a final feature-independent saliency map.

In Koch and Ullman's paper, the saliency map was a conceptual proposal with no model-based implementation (Figure 1.3A). In follow up work, the saliency based model was implemented for rapid scene analysis (Figure 1.3BC) [50]. The input to the model is a 2D image and the output is also a 2D matrix (dimensions of the matrix are the same as that of the image). The scalar values in the output matrix represent the conspicuity of each object (or pixel) in the input image. The scalar value is higher for objects or pixels that are different than surrounding objects or pixels in the input image. The first step is to extract features such as colors, intensity, and orientation in the image across multiple scales. Then a center surround difference operation is implemented on the extracted feature maps to measure how different a pixel is with respect to its neighboring pixels (Figure 1.3C). Since each feature map has its own set of units, they cannot be combined directly; they need to be normalized first. A normalization operation is used to bring all the maps to the 'same frame of reference'. Once the normalization is done, all the feature maps (across scales) are combined using a linear combination operator to obtain the saliency map (Figure 1.3B). Note that whereas each feature map has information about how salient each location is with respect to that particular feature, the final saliency map encodes for the saliency based on a combination of all features. Hence it encodes for information about the relative saliency of each location in a feature-independent manner. Once the saliency map is obtained, a winner-take-all (WTA) operation (a maximum operation) determines which location is the most salient in the input image to direct attention.

Following the identification of the most salient location, the direction of attention to it and the processing of information there, an inhibition of return (IOR) mechanism (which suppresses the salience of the currently attended location in the next iteration of computing saliency) is used to direct attention to the second most salient location. This IOR operation is an essential component of the model, because without it, in every iterative step, the same location of the image is computed as the most salient location and the viewer's attention is stuck at the same location of the scene without any update. The IOR operation in the model is motivated from psychophysical experiments in humans that showed that observers take longer to discriminate or recognize patterns in locations that they had been previously cued and attended to [51-53].

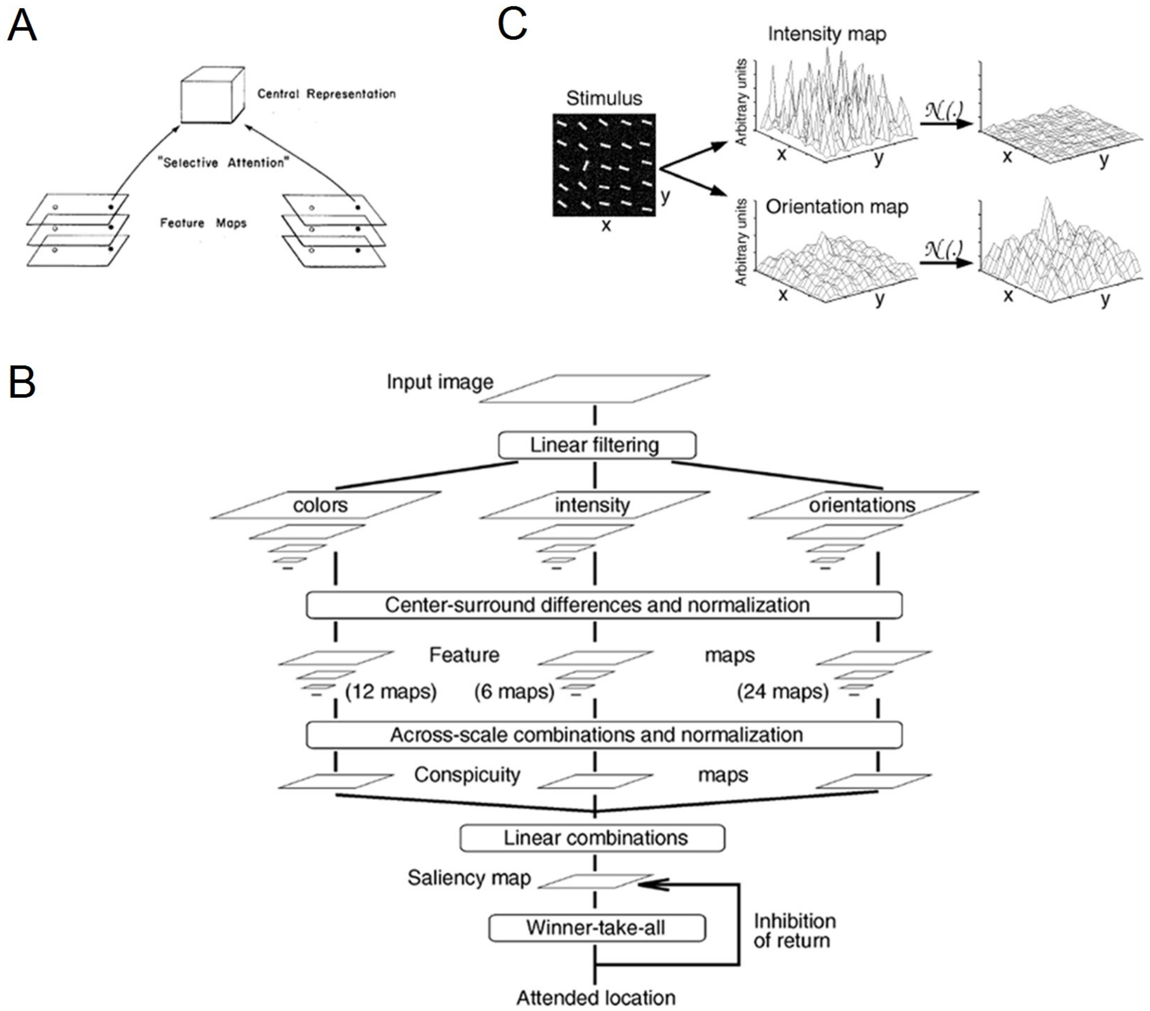


Figure 1.3: Saliency map.
(A) The original saliency map proposed by Koch and Ullman. Multiple feature maps are combined to obtain a central representation of the relative saliency of each location. (B) Saliency map as proposed in subsequent work. Feature maps are obtained from the original image across multiple scales using a center-surround difference. These maps are normalized and linearly combined to obtain the final saliency map. (C) An example showing the feature map and normalization operator on the original image. Figures adapted from [49, 50].

This step-by-step process of (a) determining the most salient location and directing attention to it, (b) inhibiting return to it and (c) directing attention to the next most salient location from the updated saliency map is iteratively applied to scan an image. The authors showed that the saliency map was able to achieve human level performance for a number of pop-out tasks and was able to select objects of interests (such as faces, buildings, vehicles) as the most important locations in natural images.

Although originally implemented to handle visual inputs (specifically, 2D images), studies have highlighted how the saliency map generalizes across sensory modalities, and can be adapted to represent auditory stimuli in the scene. One proposal was to treat the spectrogram of an auditory stimulus as the 2D image (the auditory image) from which the salience is computed [54]. However, time and frequency components of an auditory stimulus are not equivalent, since temporal information carries a lot of weight for an auditory stimulus [55]. The first model for auditory salience computation that was not based on the visual saliency map used features based on predictive coding to detect temporal saliency [56]. Whereas temporal saliency in an auditory stimulus is a key feature for directing the attention based on auditory stimuli, in the context of my dissertation, I focus on mechanisms of selection of the most important stimulus (visual or auditory) at a given instant.

1.3.3 What is a priority map?

At any instant, sensory ('external') information is not the only kind of information our brains receive (for instance, whether the traffic light was red). They must also handle information from stored memory, and information from the internal states and our voluntary goals (for instance, one's goal to turn right at a traffic light at a free-right intersection). Since salience is defined as the set of physical features of sensory stimuli, or 'bottom-up' features, that make them stand out (how bright a stimulus is, how loud a sound is, how fast an object is, etc.), the saliency map, therefore, does not include any goal driven information that the observer might have. For example, the observer might be specifically interested in searching for a person with black shirt in a crowd. In that case, the color black is 'more salient' just because of the goal driven nature of the task, even though black may not be salient, per se, as a bottom-up feature in the scene. The term used to denote the relative conspicuity of targets from a top down perspective is ‘relevance’. Priority, then, is defined as the combination of stimulus salience and stimulus relevance. A modified version of the saliency map in which goal driven biases (relevance information) modulates feature maps is called the priority map. In other words, the priority map generalizes the idea of a salience map to include the critical influence of internal goals, i.e., ‘endogenous’ or top-down information [57-59].

The original saliency (priority) map model and its numerous adaptations have proven very useful in engineering applications. They have been applied to various applications such as target tracking for moving robots [60], surveillance systems [61], predicting eye movements [62] and more.

1.4 Priority maps in the brain

Whereas ideas of priority maps arose from theoretical considerations of competitive selection for attention, and saw successful engineering applications, their original goal was to provide a framework for how brains might competitively select a target among distracters. Detailed neurobiological investigations over the years have revealed the presence of priority maps in the brain. The prerequisite properties that neurons in a brain area must exhibit for it to serve as a priority map for spatial attention are:

a. The brain area should encode space in an ordered (topographic) manner.
b. The neurons in this brain area should encode the salience/priority of each stimulus. It should not be tuned to specific stimulus features that are not intrinsically salient, such as orientation, color, sound frequency, direction of motion etc. even though it should

receive information from multiple features (much like the final saliency map receives input from the feature maps). It is also important that it receives top down information for the computation of the priority map.

c. Activating neurons (increasing their spiking rate) encoding for locations in a particular portion of the space map should increase the probability of selecting a target when it is presented in that portion of the space map.
d. Conversely, inactivating neurons encoding for locations in a particular portion of the space map should decrease the probability of selecting a target when it is presented in that portion of the space map.

One set of brain regions which immediately become the usual suspects for encoding saliency and priority are the regions in the dorsal pathway (the so-called 'where' pathway) which encode the spatial location information of the stimulus (Figure 1.4A; [63]). (The 'where' pathway is distinguished from brain areas in the ventral ('what') pathway that encode information about object identity and its details (Figure 1.4; [63]).) In the context of visual spatial attention, based on the pathway whose inputs begin in the retina and outputs end in eye movements, candidates that were thought to contribute to the implementation of priority maps were lateral geniculate nucleus (LGN), primary visual cortex (V1), visual area V4 in the extrastriate cortex, the lateral intraparietal cortex (LIP) located in the intraparietal sulcus, the pulvinar located in the thalamus, frontal eye fields (FEF) located in the frontal cortex, and the superior colliculus (SC) located in the midbrain. (Figure 1.4B; [64]). I will consider in detail, the LGN and the SC, two areas that serve, in parallel, as the first stops for visual information from the retina on the way to the rest of the brain. Additionally, these two areas illustrate a key contrast: one of them (SC) encodes a priority map, and the other (LGN) does not, thereby highlighting the neural correlates and mechanisms underlying priority maps.

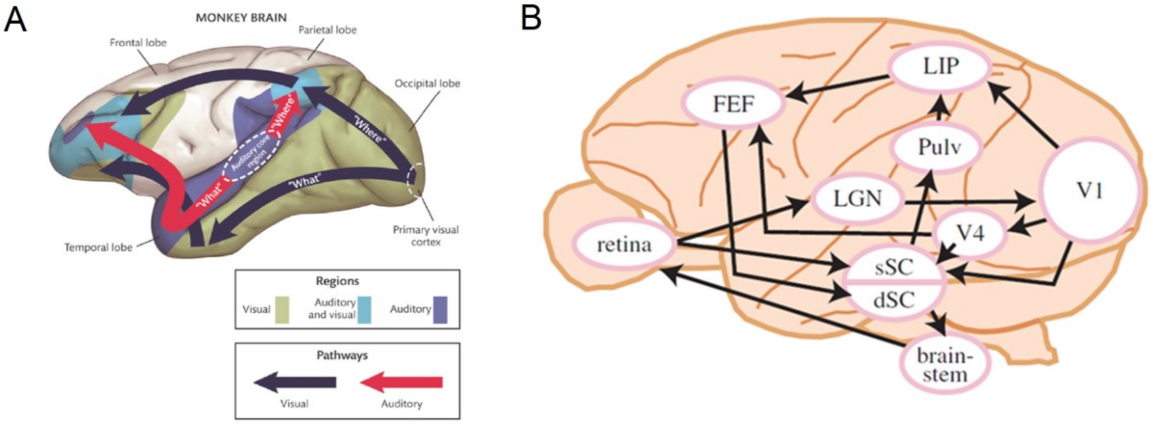


Figure 1.4: Visual pathways in the brain.
(A) The dorsal (where) and ventral (what) pathways. Figure adapted from lecture notes of Nervous system 1 offered by Dr. Stewart Hendry at JHU. (B) Visual pathways that originate from the retina input to the eye movements output. Figure adapted from [64].

1.4.1 Lateral geniculate nucleus (LGN)

In the original paper [49], the saliency map was proposed to be localized in LGN and V1, making them potentially important sites to investigate the locus saliency and priority representation in the brain. More than 80% of the axons of the ganglion axons that originate in the retina send their axons to the LGN. The other targets of the retinal ganglions axons include the hypothalamus and the superficial layers of the SC [65]. The primate LGN is a layered structure and each layer mainly receives inputs from one of three classes of ganglion cells (koniocellular (K), magnocellular (M) or parvocellular (P)). Most of the LGN neurons innervate the primary visual cortex V1. Axons from the retina terminate in the LGN in an orderly fashion resulting in a topographic map of visual space (thus meeting the key feature (a) described above). (In vertebrates, majority of the axons from the RGCs cross over the midline through the optic chiasm to the neurons in the contralateral LGN. As a result, starting from the LGN, there is a contralateral heavy representation of sensory information that carries through, until higher order areas that integrate information from both hemifields.) Even though the LGN is early in the hierarchy of the visual pathway, the visual information is well-organized, making it a potential candidate area for encoding the saliency map.

The traditional view of LGN is that of a relay station of the visual information from the retina to the visual cortex, with most of the cognitively demanding computations thought to be performed in the cortex [66, 67]. However, Francis Crick had suggested that the LGN and the collection of GABAergic inhibitory neurons around it (called thalamic reticular nucleus; TRN) might play the role of directing top down attention to the rest of the brain [68]. Till recently, due to technical limitations in recording neural data, it was difficult to test this hypothesis. However, with the advent of new techniques, recent research has shown that the LGN may play a more important role in cognitive control during demanding tasks than just relaying information to and between cortical areas. Studies in monkeys [69], humans [70], and rodents [71] have shown that LGN neurons are involved in the modulation of neural representations by attention.

However, despite these interesting findings of the role of LGN (and TRN) in attention, they can be ruled out as the sites encoding saliency or priority maps because they do not exhibit a requisite feature: LGN neurons are not insensitive to stimulus features such as luminance, contrast and color, but rather are tuned to them [72]. This results in the inability of LGN neurons to signal net saliency (uncontaminated by the encoding of specific feature values).

Based on similar arguments and experimental evidence, it is now clear that other early sensory areas such as V1/A1, V4, etc. do not encode salience maps [64]. By contrast, it is well established that the FEF, LIP, and SC encode not just salience maps, but maps of stimulus priority [73-76].

1.4.2 Superior colliculus (SC/OT): a sensorimotor hub

The superior colliculus (SC, called optic tectum, OT, in birds and other non-mammals) is a sensorimotor hub that is conserved across vertebrates. It has mutually aligned topographic maps of sensory and motor space [75, 76]. SC neurons encode the saliency and relevance of stimuli, and are not tuned to specific, intrinsically non-salient features such as color, orientation etc. Additionally, the SC has also been shown to play a critical role in the control of spatial attention. It, therefore, satisfies all of the requisite properties for a brain area to encode priority maps described at the start of section 1.4. For this reason, I focus the next few sections on covering the various studies that have highlighted the anatomical and functional properties of the SC. Since my thesis work is in barn owls, I also discuss key findings of studies in the avian homologue of the SC, i.e., the OT.

Studies in mammals and birds have shown that the SC is a layered structure. The layers can be classified into two divisions, the superficial versus intermediate-deep layers based on their anatomical connectivity and functional properties. Such a layered structure and the functional subdivisions are preserved across vertebrate species [20, 76-79]. In mammals, the superficial layers are formed by the top 3 layers and the intermediate-deep layers by the bottom 4 layers [79]. In birds, the superficial layers consist of layers 1-10, and the intermediate-deep layers, 10-15 [77]. Layer 10 is shared between the superficial and the deep layers [79].

The superficial layers of the SC/OT, referred to as SCs/OTs, receive input primarily from the retina and visual forebrain areas (striate and extrastriate areas in mammals [65, 78-81], and visual hyperpallium in birds [79, 82]). They send output primarily to 'visual' subcortical areas. In mammals, this includes the LGN and the pulvinar, which connect to the cortical areas. In birds, it includes the LGN (which connects to the visual hyperpallium) and the isthmo-optic nucleus (ION; the source of majority of the inputs from the brain back to the retina in birds) (Figure 1.5). Through these feedback projections, the SCs/OTs is able to received, send and modulate visually-related activity from and to the retina and cortical areas.

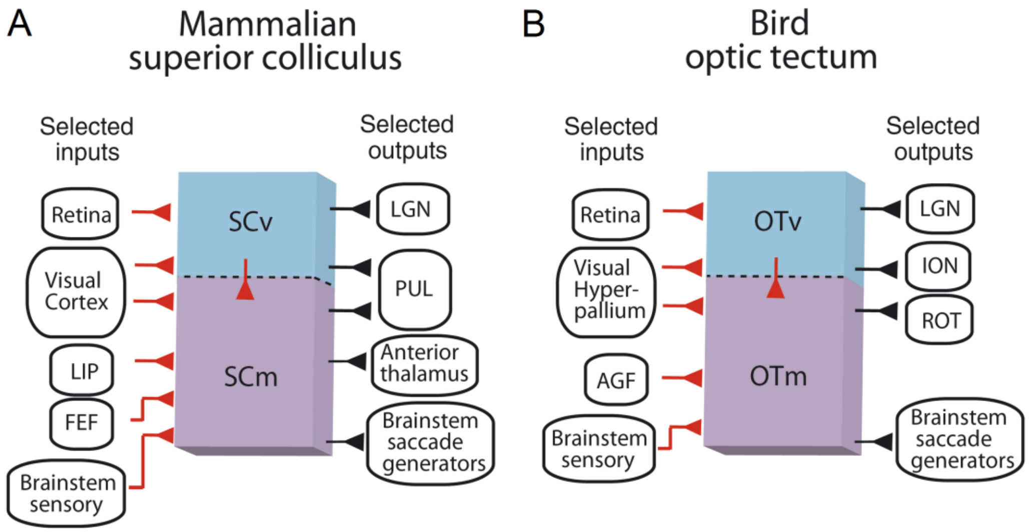


Figure 1.5: Subdivisions of the mammalian SC and the avian OT.
Input and output areas to the subdivisions are shown. The superficial areas primarily encode visual information and are anatomically connected to visual areas. The deeper layers encode multisensory and motor information and are connected to various nonstriate and subcortical areas. Figure adapted from [79].

Unlike the superficial layers, which have input-output connections primarily with visual regions in the other parts of the brain, the connections of the intermediate-deep layers provide it with more rich information. The intermediate-deep layers of the SC/OT, referred to as SCid/OTid, receive multisensory and motor-related information from the forebrain areas and the brainstem [79]. In both mammals and birds, these include the acropallial gaze fields (AGF; avian functional homologue of FEF) [83], substantia nigra pars reticulata and sensory areas of the brainstem [84-88]. Additionally, in mammals, LIP also sends information to the SCid, which sends output to the pulvinar and thalamus, which then route information to the forebrain areas. The avian OTid sends output to the nucleus rotundus (ROT; a nucleus in the avian thalamus), which connects to the forebrain. Both mammalian SCid and the avian OTid send output to the motor output areas of the brainstem [65].

1.5 Role of SCid/OTid in selection and attention

In primates and other mammals, the study of selective attention has primarily been focused on fronto-parietal brain areas (FEF and LIP), which are not as well-developed in all non-mammals. On the other hand, not only do all vertebrate brains possess the SC (called OT in non-mammals, as described above), but the SCid has been shown to play a critical role in the control of spatial attention. We note that whereas the SCid plays a dedicated role in directing spatial attention, fronto-parietal brain areas are involved in other forms of selective attention as well (feature-based, object-based, non-spatial), and in other cognitive tasks.

1.5.1 Sufficiency (activation studies)

Contrary to traditional thinking that selective spatial attention is controlled exclusively by fronto-parietal networks, numerous recent studies have revealed an equally critical role for the SCid in spatial stimulus selection and spatial attention.

Studies have shown that focally activating the SCid (increasing the activity of neurons within a small patch of SCid tissue) results in spatially specific selection enhancements. Subthreshold electrical microstimulation of SCid neurons increases the probability of selecting or pursuing (following with eye movements) a target presented inside the spatial RFs of the stimulated SCid neurons [89]. The microstimulation is 'weak' (or sub-threshold) because it does not, by itself, induce spikes. Rather, it serves as bias to increase spiking activity of neurons by bringing their membrane potential closer to spiking threshold. In one study, authors showed that activation of SCid neurons could act in place of a spatial cue to bias attention towards a particular location. In addition, weak stimulation of a portion of the SCid space map improved the ability of monkeys to detect a small change in the information content of a target stimulus presented inside the RF of the stimulated neurons [90]. No such improvements were found when the target was outside the RFs of the stimulated neurons. Similar results were also found in another study where monkeys performed covert attention task [91]. This suggested that the SCid might also play a role in covert attention, and that the execution of covert and overt attention might share a common neural mechanism - a group of SCid neurons called visuomotor neurons. These studies together provide convincing evidence that activating SCid neurons increases the probability of target selection specifically when it is located inside the region encoded by those neurons.

1.5.2 Necessity (inactivation studies)

Whereas the above studies have shown that activating SCid neurons increases the probability of selecting a target, thereby establishing SCid's sufficiency for attention control, studies involving silencing, or 'inactivating', the SCid have demonstrated the necessary role the SC plays in spatial attention.

A 2004 study in monkeys investigated the role of SC in saccade target selection by focally suppressing the activity of SCid neurons (using lidocaine, a sodium channel blocker or muscimol, a GABA agonist) [92]. Monkeys performed a color-oddball task by making a saccade to the one stimulus with a color different from that of three other stimuli of the same color. The monkey made significantly higher errors when the target was located inside the RFs of the inactivated SCid neurons. The monkey's ability to saccade to the target was unaffected when the target was presented inside the RF of neurons in the unaffected parts of the SCid (Figure 1.6). Additionally, such a deficit was only present when the monkey had to select the target in the presence of distractors. Its ability to make saccades to a single stimulus presented inside the inactivated region was unaffected. Therefore, the performance deficit was not due to an inability to 'see' the stimulus, or to execute eye movements to it, but rather due to deficits in selecting a target in the presence of distractors. This role of SCid was shown to be especially critical when the difficulty of the task was high, further demonstrating that the SCid played a role in overt target selection.

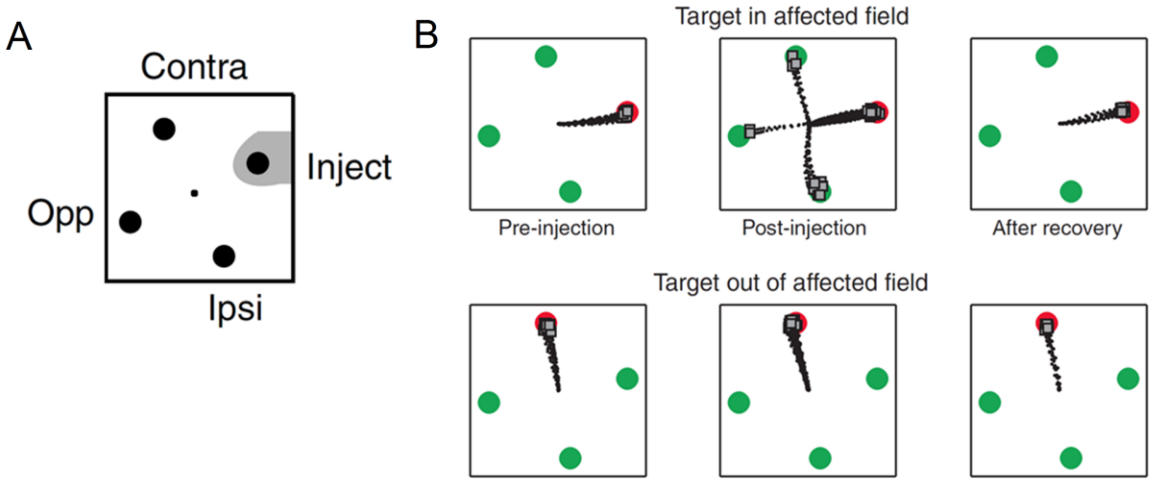


Figure 1.6: SCid is required for overt target selection.
(A) Monkeys performed a saccade target selection task where they selected the odd colored target in the presence of distractors, and made a saccade to it. A portion of the SC space map corresponding to one of the four locations was focally inactivated. (B) When the target was located in the affected portion, the monkey's performance accuracy decreased (top 3 panels). Performance remained unchanged when the target was outside the affected field (bottom 3 panels). Figure adapted from [92].

Convincing evidence that SCid is also required for *covertly* selecting a target was presented in 2010 [93]. In this task, monkeys were required to fixate on a central dot throughout the trial and cued to attend to one of four locations using a color ring. Next, stimuli consisting of stochastically moving dots appeared briefly in each of the four locations. The cue then disappeared. Following a delay, the dots in the cued location and the location diametrically opposite to the cued location exhibited coherent directional motion. Then the stimuli disappeared and 4 response boxes appeared. The task for the monkey was to make an eye movement to the response box corresponding to the direction of coherent motion of the dots in the cued location (while ignoring the foil stimulus' direction). The clever design of the task allowed the authors to dissociate the location of covert attention (which depended on which location was cued) and the location of eye movements (which depended on which direction the cued stimulus moved). The monkeys performed very well (~75% correct responses) on this fairly complicated task (Figure 1.7A).

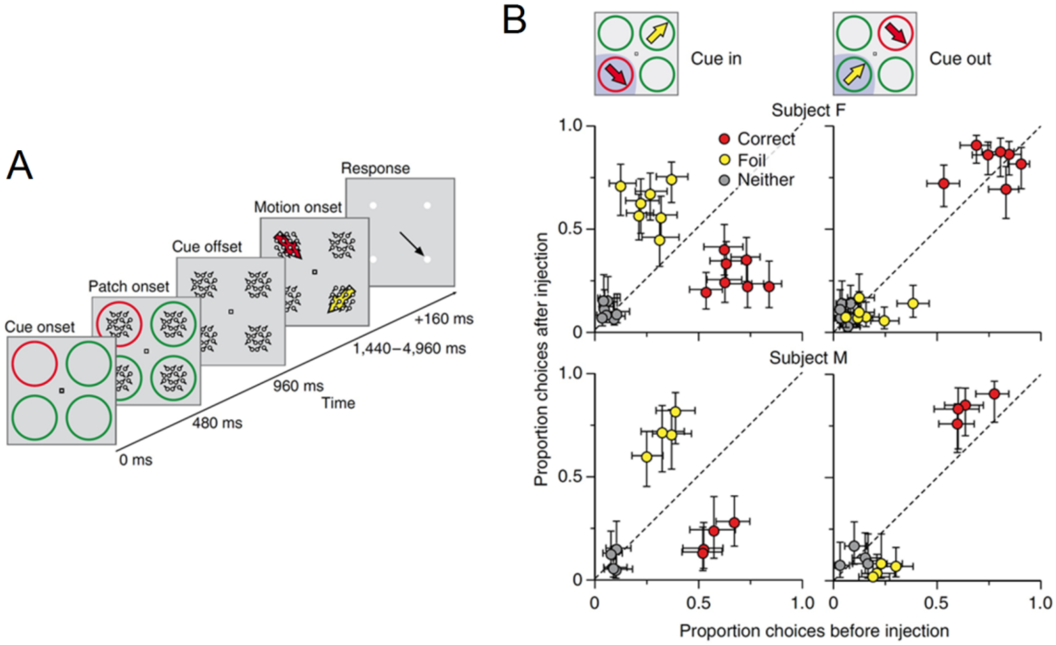


Figure 1.7: Scid is required for covert target selection.
(A) The task. Monkeys were cued to attend to one of four locations and respond with eye movements based on the direction of motion of dots at the cued location. (B) Performance. The performance of the monkeys was worsened if the cued location was affected. Monkeys made more eye movements based on the foil stimulus and ignored the cued stimulus when the cue fell inside the affected region (left two panels). Performance remained unchanged when the target was outside the affected field (right 2 panels). Figure adapted from [93].

Next, the authors inactivated portions of the SCid space map corresponding to one of the four locations using muscimol, and discovered very similar effects in this covert attention task as in the previous study of overt target selection. The performance of the monkeys was degraded when the cued location was inside the RFs of the inactivated SCid neurons (Figure 1.7B). Additionally, even though there were minor motor deficits in the affected region, it could not explain the behavioral deficits in target selection. Indeed, such target selection deficits were also found in a version of the task that used button-press instead of eye movements as responses to the task. This was the first study demonstrating that SCid was required for covert target selection.

The above microstimulation and silencing studies have demonstrated a causal role for the SCid in target selection in both covert and overt attention tasks. But what are the computations that take place in the SCid to implement target selection? Work over the past 10 years in birds, specifically, barn owls, but also pigeons and chicks, have started to yield insights.

1.6 Barn owl midbrain attention network

The midbrain attention network is composed of two components. One component is the multilayered optic tectum (SC/OT). The other is the isthmic complex satellite to the SC. It includes two specialized satellite isthmic nuclei - the GABAergic isthmi pars magnocellularis (Imc), and the cholinergic isthmi pars parvocellularis (Ipc).

The OT is a 15-layered structure in birds, and has been shown, across the layers, to contain topographic maps of visual space (layers 1-10), and multisensory and motor space (layers 11-15), for locations in the contralateral hemifield. Specifically, spatial azimuth is encoded along the rostrocaudal axis of the OT, and spatial elevation is encoded along the mediolateral axis (Figure 1.8A-C). The sensory and motor maps are mutually aligned. (I will discuss the isthmic complex starting section 1.8).

Barn owls are multisensory specialists, possessing high acuity visual as well as auditory systems, making it possible to study spatial selection and attention across different sensory modalities and potentially eludicate circuit mechanisms that generalize and are sensory-modality independent. Additionally, a biological feature of barn owls is that their eyes and ears are virtually immobile in their heads. Due to this feature, the visual and auditory maps of a barn owl are always aligned. Therefore, an experimental setup where the head is immobile virtually fixes the topographic maps of visual and auditory space allowing reliable multisensory measurements from specific portions of the space map that are not confounded by eye movements. These two characteristics of barn owls make them a particularly useful model to study the neural basis of spatial computations.

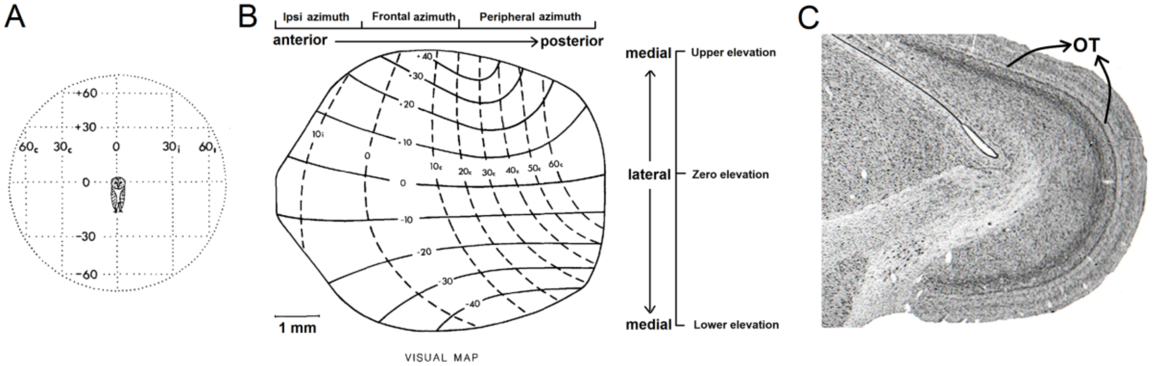


Figure 1.8: Space coding in the barn owl.
(A) Azimuth and elevational encoding of space using double pole coordinate system. (B) The topographic map of the barn owl midbrain. Azimuth is encoded systematically along the rostrocaudal axis and elevation is encoded along the mediolateral axis. +ve; upper elevational locations; -ve- lower elevational locations; i-ipsilateral space; c-contralateral space. (C) The inverted C shaped structure are the layers of the OT.

1.7 Correlates of stimulus selection across space in the barn owl OT

The presence of a precisely aligned topographic multisensory and motor space map in the OTid makes it an ideal site in the brain to study the mechanisms of stimulus selection. A series of studies in birds in the past decade have revealed fundamental mechanisms that the midbrain uses to solve spatial stimulus competition and signal the highest priority location.

A study in 2010 in head fixed barn owls revealed the mechanisms of global spatial stimulus selection in the OTid space map [75]. Neurons in the OTid were found to encode the salience of a stimulus in a featureless manner. A distant competing stimulus presented outside the RF of an OTid neuron was shown to suppress the responses of the OTid neuron to a stimulus presented inside its RF. The competitive inhibition operated in a modality independent manner (i.e., it didn't matter if the competing stimulus was visual or auditory). The inhibition was long-range and extended well beyond the traditional center-surround inhibition (a competing stimulus well beyond 60˚ was able to suppress the responses of the OTid neuron). More interestingly, it was found that the strength of inhibition was strong nor matter the location of the competing stimulus as long as it was outside the OTid neuron's RF (Figure 1.9AB). In other words, the competitive inhibition that mediates stimulus selection in the OTid space map was essentially location-invariant, allowing for the selection of the highest priority stimulus no matter where it occurred in the OTid space map. (A major part of my dissertation (Chapters 2 and 3) focuses on the investigating the circuit mechanisms underlying the implementation of such location-invariant stimulus selection in the OTid space map.)

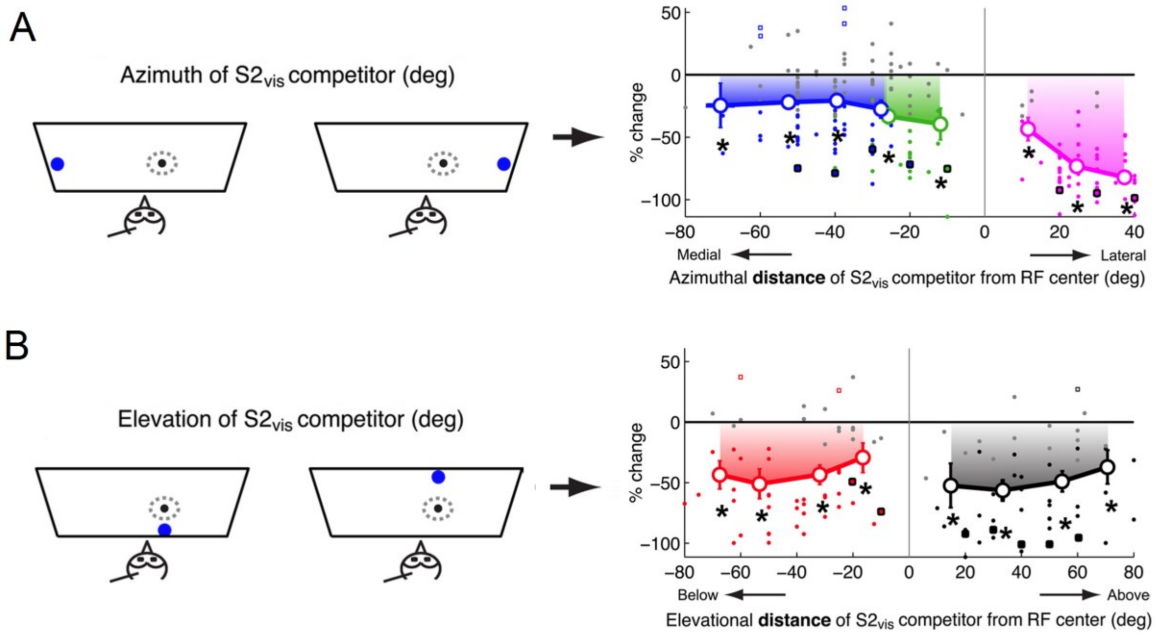


Figure 1.9: Global competition in the OTid along.
(A) azimuth, and (B) elevation. Left panels: Stimulus protocol used to measure the effect of the location of a stronger competing stimulus. Right panels: Strength of inhibition due to the competing stimulus as a function of distance from RF center. The strength of inhibition is strong irrespective of the actual location of the competing stimulus. In other words, the stimulus selection in the OTid operates in a location-invariant manner. Figure adapted from [75].

A follow up study characterized the nature of inhibition due to a competing stimulus by varying its strength [94]. The authors found that the responses of the OTid neuron decreased as the strength of the competing stimulus increased (Figure 1.10A-C). Of specific interest were a subset of OTid neurons that exhibited 'switch-like responses (Figure 1.10C) (the remaining OTid neurons exhibited gradual responses). The switch-like responses had a high firing rate when the RF stimulus was stronger than the competing stimulus, and 'switched' to lower firing rates when the competing stimulus became stronger than the RF stimulus (Figure 1.10E). About 30% of the OTid neurons exhibit switch-like responses (Figure 1.10D). Additionally, the switch-value in the OTid responses depends on (and is roughly equal to) the strength of the RF stimulus (Figure 1.10E-G). This results in a flexible, and relative stimulus strength dependent representation of responses in the OTid.

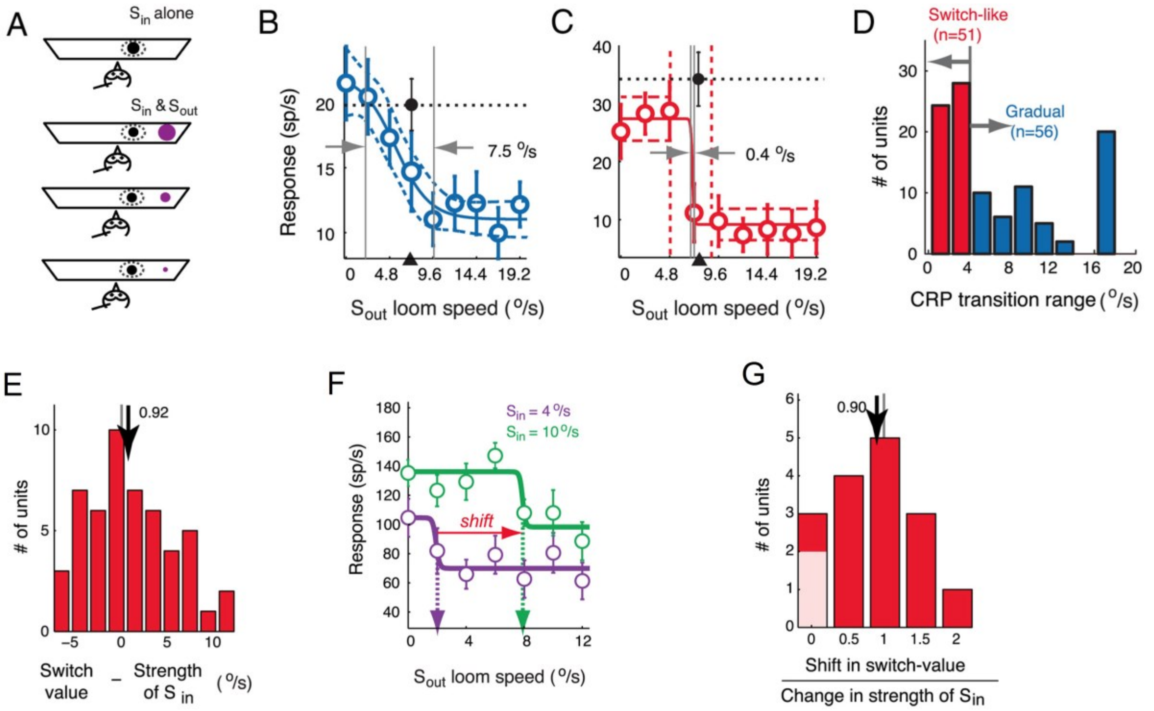


Figure 1.10: Switch-like and gradual neurons in the OTid.
(A) Protocol for studying stimulus competition in the OTid. One stimulus of fixed strength was located inside the RF of an OTid neuron and a second stimulus was presented outside the RF. The strength of the competing stimulus was systematically varied while recording the responses of the OTid neuron. (B-C) Responses of the OTid neuron plotted as a function of the competing stimulus strength showing. Responses can either be (B) gradual or (C) switch-like. (D) The number of switch like neurons are roughly equal to 30% of the number of gradual neurons. (E) The switch-value of the OTid neuron depends on the strength of the RF stimulus. The responses of the OTid neuron switches from high to low when the strength of the competing stimulus exceeds the strength of the RF stimulus. (F) Response of an OTid neuron to two different values of RF stimulus strength. The switch-value shifts when the RF stimulus becomes stronger by the same amount as the change in the RF stimulus strength (G) Summary of the shift in switch-value when the RF stimulus strength is shifted. They are roughly equal resulting in flexible relative-stimulus strength dependent stimulus selection. Figure adapted from [94].

The role of switch-like neurons is of specific interest in stimulus selection because of the advantages they yield in comparison to the responses of gradual responses. For instance, switch-like neurons result in higher discriminability in neural responses when the competing stimuli are close in strength [94]: As long as the stronger of the two stimuli is inside the RF of the OTid neuron, it fires at a high rate (even if the competing stimulus is close in strength,). This results in a (flexibly) categorical representation of the competing stimuli across the OTid space map, with the two categories being $s_{in} > s_{out}$ and $s_{in} < s_{out}$ [95]. In other words, the OTid signals of the highest priority stimulus in a winner-take-all-like manner, a signal that is sent to the rest of the brain and the motor areas to guide attention or eye movements as required. Such switch-like responses can also explain the behavioral deficits (or improvements) seen in monkeys when neurons in the SCid space map are focally inactivated (or stimulated) [92] (Figure 1.6 and Figure 1.7).

1.8 Imc: GABAergic mechanism for stimulus selection

It is clear that response inhibition evoked by a competing stimulus is an essential part of generating switch-like categorical responses (or relative stimulus strength dependent inhibition more generally). However, the necessary long-range inhibition cannot come from within the SC: it has been shown that the SC only has short range, classical surround inhibition [96, 97]. The source of the inhibition turned out to lie in the other component of the midbrain attention network, namely the isthmic complex.

The SC interconnects heavily with the two nuclei in the isthmic complex. One component of the GABAergic isthmic nucleus called the isthmi pars magnocellularis (Imc). The other component is a cholinergic isthmic nucleus called as the isthmi pars parvocellularis (Ipc) (Figure 1.11A). The isthmic complex also hosts two other nuclei called the nucleus semilunaris (SLu) and disseminated nucleus (ID). In mammals the Ipc and SLu together form the parabigeminal nucleus. The role of the ID in stimulus selection is relatively unknown, and for this reason I do not discuss it further (The ID is thought to be important for cross hemispherical competition, but this is yet to be demonstrated.) In my dissertation, the functional logic of Imc (and to an extent the Ipc), along with the OT in service of stimulus selection, forms the focus of the investigations.

The components of the midbrain attention network are conserved across all vertebrates [79, 85-87, 98]. However, they are anatomically well separated in birds making recordings from different components of the network easier in birds. Hence they have primarily been studied in birds (barns owls, pigeons and chickens).

The isthmi pars magnocellularis (Imc) is a GABAergic nucleus in the midbrain attention network. Its functional role was first described in turtles based on the anatomical connectivity with the OT [99]. In a study in chicks, Imc neurons were shown to receive focal inputs from neurons in layer 10 of OT (OT10). Imc neurons send back broad inhibition to OTid but specifically spare the portion of the space map from which they receive input resulting in a 'donut-like' anatomical connectivity (Figure 1.11B).

Like the Imc, the Ipc neurons also receive focal inputs from neurons in OT10. However, unlike Imc, they send back focal amplification to neurons across all layers of the OT in the same portion of the space map that they receive input from (Figure 1.11B). Additionally, Imc also sends broad range inhibition to the Ipc. However, whether Imc axons spare a portion of the Ipc, like they do in OTid is unknown (Figure 1.11B). I explore this question and the role that it plays in stimulus selection in parts of Chapter 4 of this dissertation.

The pattern of anatomical connectivity in the midbrain attention network is well suited to support stimulus selection in the OTid space map. The broad range inhibition from Imc is thought to facilitate distractor suppression, and the focal Ipc amplification is thought to amplify the RF target stimulus resulting in relative-stimulus strength dependent response profiles. Together the Imc and Ipc connectivity can explain the results found in previous studies (Figure 1.6 and Figure 1.7).

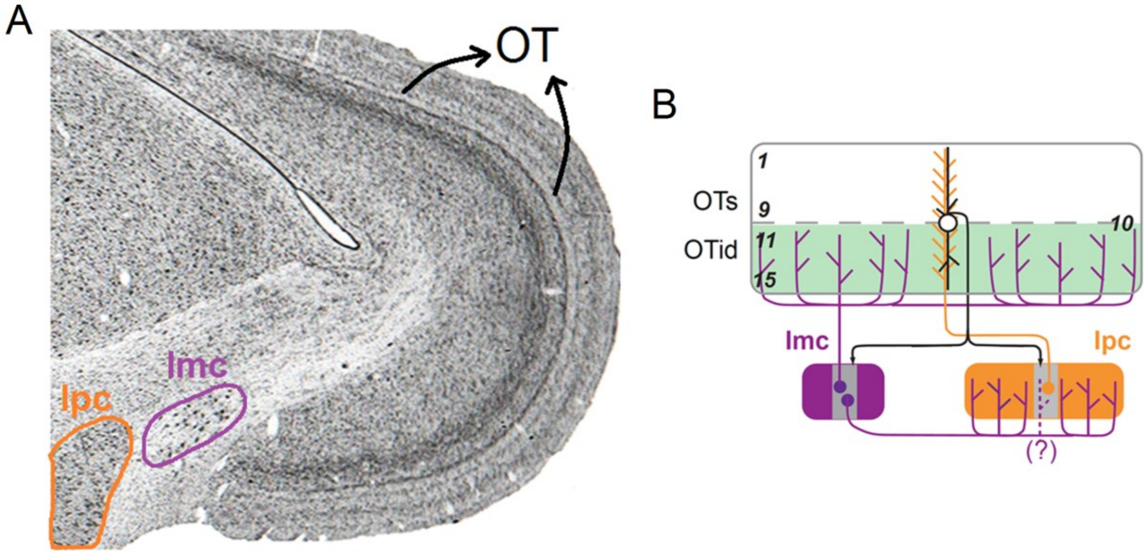


Figure 1.11: Anatomical connectivity in the midbrain attention network.
(A) Coronal section showing the components of the midbrain attention network. (B) Imc neurons receive focal input from OT neurons in layer 10. They send broad inhibition to OTid and Ipc. In the OTid, Imc spares inhibition to the portion of the space map they receive input from. Whether similar sparing of inhibition by Imc is also present in the Ipc is unknown. Ipc neurons receive focal input from neurons in layer 10 of the OT space map. They send back strong focal amplification to all layers of OT in the same portion of the space map.

The role of Imc as a source of competitive inhibition was first suggested based on its anatomical connectivity [99]. It was directly tested in a study that focally inactivated Imc neurons during stimulus selection, while simultaneously recording from OTid neurons to study stimulus selection [100]. As seen previously, the presence of a stronger competing stimulus outside the RF results in suppression of OTid responses to a stimulus presented inside its RF. This suppression is abolished when the Imc neuron encoding for the competing stimulus is focally inactivated using iontophoresis (with kynurenic acid, a pan-glutamate receptor blocker). The competitive inhibition returns to a strong level upon the recovery of the Imc activity, suggesting that the competing stimulus sends inhibition to the OTid neuron encoding for the target (Figure 1.12A-D).

More interestingly, Imc is also the source of the 'switch-like' responses in the OTid. Inactivating Imc abolishes relative stimulus strength dependent switch-like responses in the OTid. So the ability of the OTid to signal the strongest stimulus in a winner-take-all manner is a result of inhibition from the Imc (Figure 1.12E-G).

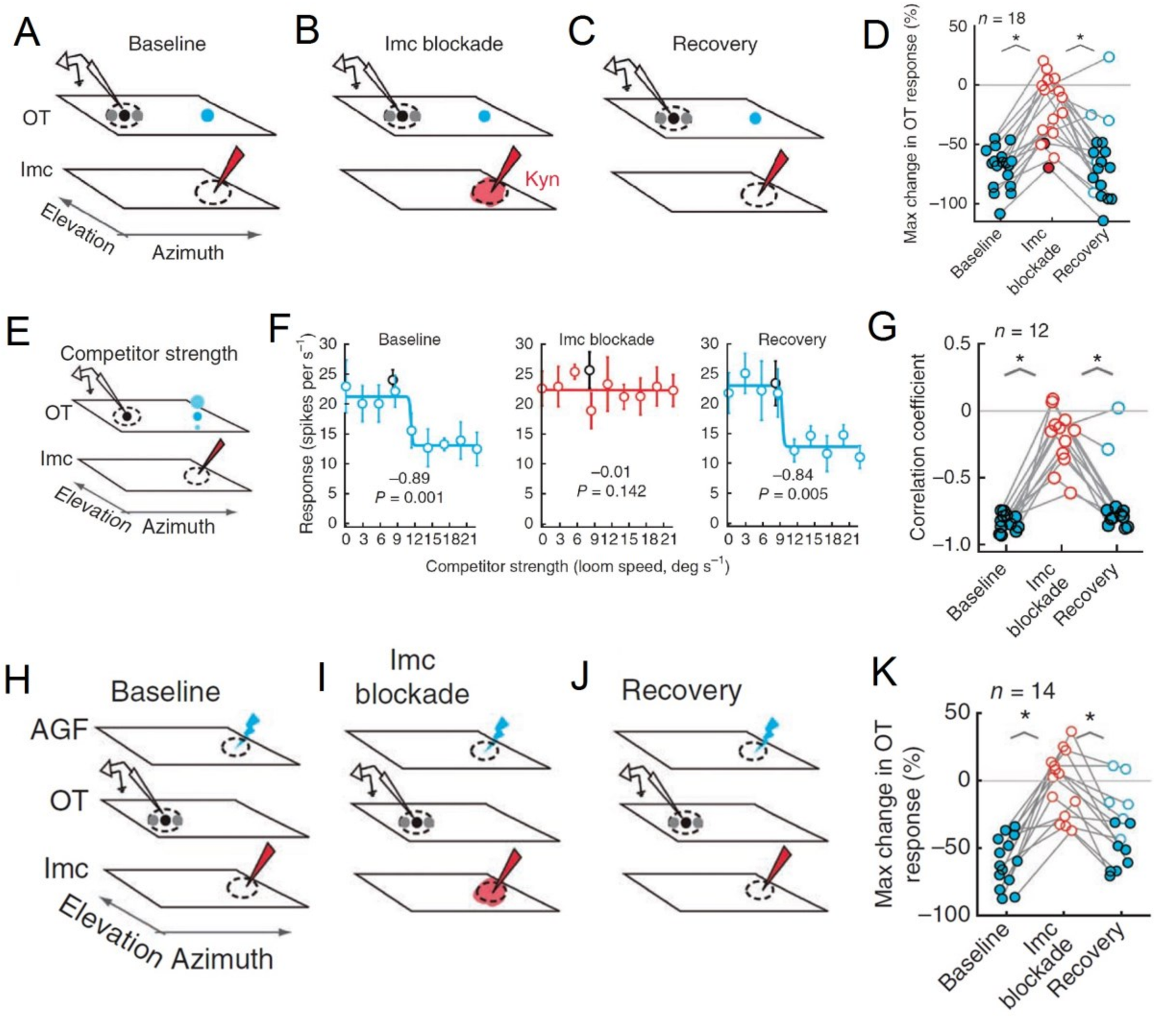


Figure 1.12: Role of Imc in endogenous and exogenous stimulus selection in the OTid space map.
(A-C) One stimulus is presented inside the RF of an OTid neuron, and a competing stimulus is presented at a distant location. Activity is recorded from the OTid and simultaneously the Imc neuron encoding for the competing stimulus when Imc is (A) intact, (B) inactivated and (C) recovered from inactivation. (D) Inactivation of Imc abolishes response inhibition driven by the competing stimulus. (E) Stimulus protocol used to measure the effect of stimulus strength on the response of OTid neuron. (F) The OTid neuron has a switch like response (left panel), which gets abolished when Imc is inactivated (middle panel), and becomes switch-like again when Imc activity is recovered. (G)

Summary showing that Imc is the source of switch-like (and gradual) responses in the OTid. (H) Stimulus protocol to study the role of Imc in endogenous stimulus selection. A stimulus is presented inside the RF of the OTid neuron and a distant AGF neuron is microstimulated, while simultaneously recording from an Imc neuron when its activity is (H) intact, (I) inactivated, and (J) recovered from inactivation. (K) Summary showing that AGF microstimulation results in endogenously driven response suppression at a distant OTid site, which gets abolished when Imc is inactivated, and returns when Imc activity is recovered from inactivation. Figure adapted from [100].

Additionally, Imc is also required to mediate top-down (exogenous) stimulus selection in the OTid space map. The top down bias signals enters the OT through the AGF in a topographic manner. Microstimulation of AGF neurons (the AGF is the functional homologue of the mammalian FEF [83]) using sub-saccadic currents is shown to cause space specific response gain at the aligned OTid RFs by covertly shifting the attention to the stimulated region consistent with studies in monkeys [101, 102]. It also results in response suppression of OTid neurons with non-aligned RFs resulting in endogenous competitive inhibition in the OTid [101]. When Imc neurons aligned with the AGF neurons were inactivated, it resulted in the abolishing of endogenously driven inhibition at distant OTid sites. The endogenous inhibition returned when the Imc activity was recovered showing a causal role for Imc in endogenously driven competitive stimulus selection (Figure 1.12H-K).

In addition to suppressing the OTid and Ipc neurons, Imc neurons also suppress each other. Modeling efforts have suggested that such reciprocal inhibition between inhibitory neurons is the most efficient circuit motif to achieve flexibility, an essential component of stimulus selection (Figure 1.10F). Using in-vitro slice experiments, it was shown that the Imc neurons do in fact inhibit each other and have this efficient circuit motif. Whether this circuit motif is in fact required for flexibility in the OTid during in-vivo experiments of stimulus selection is still unknown.

The Imc plays an indispensable role in mediating competitive interactions through multiple ways. It drives both endogenously and exogenously driven competitive interactions in the OTid. The Imc sculpts relative stimulus-strength dependent 'switch-like' activity in the topographic OTid space map, which is essential for selecting the most important location to guide attention in a covert and overt manner. This essential role played by the Imc in stimulus selection makes it an extremely crucial part of not only the midbrain attention network, but also more broadly the oculomotor network.

1.9 Ipc: cholinergic enhancement of selected stimulus

The Ipc receives focal inputs from the OT layer 10, and sends output to the same portion of the space map across all layers of the OT. Majority of its axons terminate in the superficial layers of the OT (which receive visual inputs from the retina and the visual hyperpallium) [86]. Like the OTid, the Ipc also has aligned visual and auditory space maps [103] that encode for stimuli in a relative stimulus strength dependent manner [104]. Like the OTid, a fraction of the Ipc neurons (30%) also exhibit switch-like responses. The long range stimulus competition in the Ipc is mediated by Imc neurons [87]. Since the Ipc also encodes for space in a stimulus-strength dependent manner, the output that Ipc sends across the OTid layers only amplify the strongest stimulus.

Despite playing an important role in enhancing the representation of the target, the ability of the Ipc to do so critically depends on it representation of relative stimulus strength, which is ultimately driven by Imc. Hence, the Imc plays an indispensable role and is necessary for stimulus selection in the midbrain network.

1.10 Scientific contributions of this thesis

In this introductory chapter, I have described selective attention as a mechanism to overcome the information bottleneck. I then presented a modeling framework using the saliency and priority map model for thinking about how selection of the target of attention might occur. In this context, I have pointed out the insights available thus far regarding how selection may be implemented in neural circuits, and these have come from the study of spatial selection (in the broader context of spatial attention) in the barn owl (and other birds). These experiments involved investigations of a specific midbrain network that is critically involved in the control of attention, namely, superior colliculus and its satellite nuclei. However, several fundamental questions remain open regarding the neural implementation in this network of computations that are essential for the selection of the most important stimulus across space.

In this dissertation, we present our findings on the neural circuit mechanisms underlying two essential computations of stimulus selection. Specifically, with experiments, theory and computational modeling studies, we have discovered several fundamental insights regarding how the Imc-OT network implements these computations in its circuitry. This work has broad implications to the neurobiological basis of selection (beyond just attention) as well as for engineering applications.

The nature of encoding of space by Imc is still a relatively unanswered question. This is a well answered in a detailed manner in the OT, but an understanding of how the Imc encodes space is essential to deconstruct the role it plays in stimulus selection. In Chapter 2, we present our discovery of an unusual form of encoding of visual space in the Imc. We discovered that Imc neurons encode visual space using unusual, multi-lobed receptive fields. Such receptive fields have not been reported in visual sensory areas prior to our findings. Our study is the first such report to the best of our knowledge. Further, we characterize the topographical properties of Imc along azimuthal and elevational space in the face of multi-lobed encoding of space. Finally, we present and validate the hypothesis that multilobed RFs are necessitated by the scarcity of Imc neurons in the context of the challenging problem it is solving i.e. location-invariant selection.

In Chapter 3, we present a model that is grounded in the anatomical connectivity of the midbrain selection network, and also incorporate the scarcity of Imc neurons. The model reveals how Imc solves one of the computations essential for selection, namely location-invariance, the ability to select the more salient among two competing stimuli independent of the actual locations the stimuli occupy in the world. We then demonstrate that the key properties predicated by the model hold true in the barn owl midbrain selection network.

The Imc is known to be the source of switch-like categorical representation in the OTid space map. But how is the Imc able to result in such switch like responses? Which aspect of the circuitry or the functional mechanisms of Imc are essential for such switch-like responses? This is an extremely important question not just for answering stimulus selection in the OTid space map, but also to understand a more general computational question across other behavioral tasks and activity in the brain. In chapter 4, we investigate the neural circuit mechanisms generating categorical representations of competing stimuli (and of robustness to sensory and neural noise). Using modeling of the midbrain circuit and with extracellular recordings, we show that the midbrain uses an exquisitely structured donut-like pattern of net functional connectivity to achieve such categorical responses. Additionally, the donut-like connectivity is 'multi-holed' consistent with the role of Imc in combinatorial inhibition for stimulus selection across space. We further compare donut-like pattern of connectivity to other motifs that have been hypothesized to play a role in categorical representation of stimuli.

2. Encoding of space by the Imc

In this chapter, I present our findings on the visuospatial encoding properties of Imc neurons (Figure 2.1A). An understanding and a thorough characterization of the spatial encoding properties of the Imc is essential in the context of the indispensable role it plays in endogenous and exogenous stimulus selection (Figure 1.12, [100]).

Mechanistically, competitive suppression in the OTid is orchestrated by the GABAergic Imc through its specialized anatomical connectivity with the OT [85, 87, 100]. Neurons in layer 10 of the OT (OT10) provide input to each Imc neuron, which, in turn, projects back broadly across the OTid space map except to the portions that encode the input locations [85] (Figure 2.1B). This anatomy allows the Imc to implement a spatial inverse operation, distributing priority-dependent inhibition to all competing locations in the OTid space map (Figure 2.1C). Notably, inactivation of the Imc abolishes this competitive inhibition as well as spatial selection in the OTid (Figure 1.12; [87, 100]).

In this context, if the spatial RFs of Imc neurons are assumed to be small, resembling those of the input OT10 neurons, and possessing the same topographic property as them, a conceptually straightforward strategy by which the Imc might achieve selection for all possible pairs of locations in the OTid is illustrated in Figure 2.1D. For any given pair of stimulus locations, each stimulus in the pair would activate a group of neighboring Imc neurons encoding for adjacent locations, resulting in inhibition with a spatial pattern that would suppress the neurons encoding the other stimulus (Figure 2.1D). Simply repeating this Imc-OT circuit module for all location-pairs would successfully implement stimulus selection across space with a strategy termed 'modular copy-and-paste'.

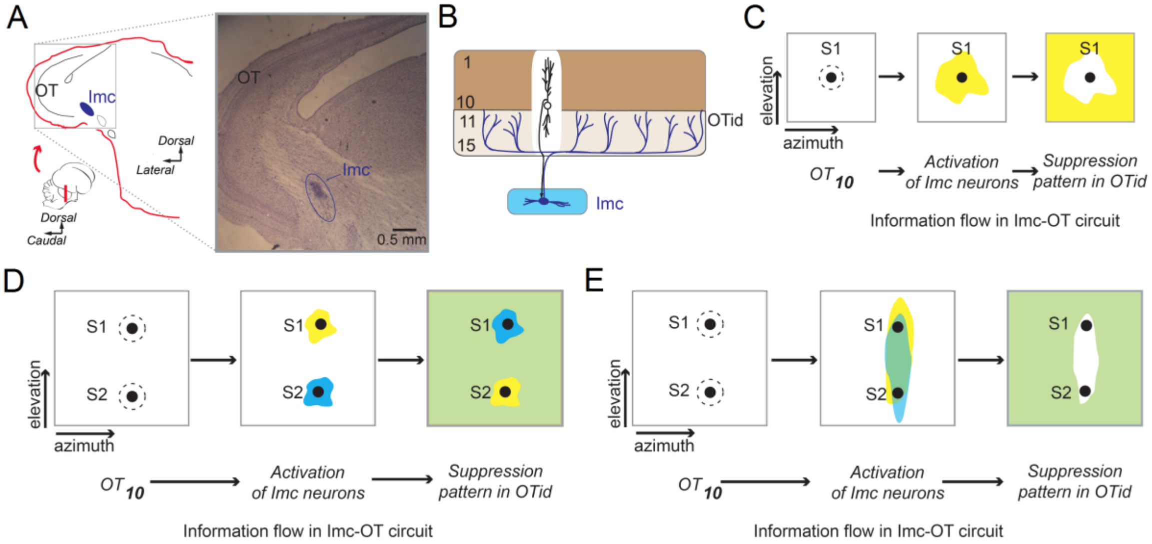


Figure 2.1: Anatomical connectivity and information flow betweeen the Imc and OT. (A) *Left:* Cartoon showing side view of barn owl brain (inset), and coronal section taken along the indicated line in inset. *Right*: Nissl stained, coronal section of midbrain depicting the multilayered optic tectum (OT) and the isthmi pars magnocellularis (Imc). The OT10 is seen as a darkly stained arc of cell bodies. The Imc is a long and narrow (baguette-like) structure: 2800 μm long rostrocaudally and 350 μm mediolaterally; appears in transverse sections as a 700-μm x 350-μm elliptical disk of neurons (blue oval) [100]. The long, rostrocaudal axis of the Imc is parallel to the rostrocaudal axis of the OT. Dark area in the dorsal portion of Imc: electrolytic lesion following stereotactic and electrophysiologically-based targeting of Imc. (B) Schematic of anatomical connectivity between the Imc and OT. Imc neurons are known to receive input from a focal portion of OT10 in azimuth and a large portion of OT10 in elevation. In turn, they project broadly back to the OTid sparing just the portions of OT providing input [85]. Shown is an example OT10 neuron (in black) that provides input to the Imc neuron (in blue). Projections from this Imc neuron back to OTid are widespread (blue lines) while sparing the input neuron (white region, across OTid layers). All layers of OT are known to represent space topographically, but how the Imc represents space is not well understood (see also (E)). (C) Schematic of information flow

through the OT10- Imc-OTid circuit showing the functional, spatial-inverse operation executed by established Imc-OT connectivity [85]. Maps of visual space in the OT10 (*left*), Imc (*middle*) and OTid (*right*). For purposes of illustrating the spatial inverse operation, Imc RFs are assumed to be large with an unknown shape (yellow shading). A visual stimulus S1 at location 1 activates the space map in OT10 (left; dashed circle - RF of activated neuron). This, in turn, activates an Imc neuron (middle; yellow represents assumed RF of activated Imc neuron), which delivers inhibition to all locations in the OTid space map that are outside the RF of the activated Imc neuron (right; yellow shading) [85, 87, 100]. (D) Schematic representation of stimulus selection in the OTid under the assumption that Imc RFs are small, resembling OT10 RFs. *Left*: Shown are two stimuli S1 and S2, at locations 1 and 2, respectively, which activate corresponding neurons in the OT10 space map. *Middle*: Imc neuron activated by S1 (yellow RF), and Imc neuron activated by S2 (blue RF). *Right*: Combined pattern of suppression generated in the OTid by the activated Imc neurons: each neuron delivers suppression to locations outside its RF; green = yellow + blue. Each stimulus successfully suppresses the other – S2 lies within the yellow zone of suppression produced by S1, and vice-versa – implementing selection for stimuli at these two locations. Similarly, if every spatial location was encoded by an Imc neuron with a small, OT10 like RF (and with space-inverting connectivity with the OT), then stimulus selection in the OTid would be achieved successfully for all pairs of locations (the 'modular copy-and-paste' strategy described in the text). (E) Same as (D), but with Imc RFs that are large and elongated vertically, covering almost the entire elevational extent, as reported in the literature [105, 106]. Shown in the middle panel are the RFs of two Imc neurons, in yellow and blue, respectively. *Left*: As in (D). *Middle*: S1 activates both Imc neurons, and so does S2. *Right*: Resulting patterns of inhibition in the OTid space map; green = yellow + blue; large swaths of space are left without inhibition (white region in right panel, corresponds to intersection of the two RFs). Specifically, neither stimulus is suppressed by the other even though the two stimuli are well separated in elevation (shown here to be approximately 80° apart), preventing stimulus selection. In fact, with such elongated RFs, stimulus selection cannot be solved for any the elevational-pairs that lie in the intersection region of the elongated Imc RFs −approximately 40% of the total number of elevational-pairs at this azimuth. Large, vertically elongated Imc RFs, therefore, are unable to support spatial selection in the OTid across all elevational locations. This is an apparent paradox in terms of Imc-OT function because the OTid is known to solve selection, including when stimuli are < 10° apart in azimuth or elevation [75, 93, 94, 107], with Imc driving this global competitive selection [100].

Despite the indispensable role Imc plays in stimulus selection, the precise nature of the spatial RFs of Imc neurons, and their properties are not well understood. Previous work in the Imc suggests that Imc neurons have spatial RFs that are vertically elongated, covering almost the entire extent of elevational space in an uninterrupted manner [105, 106]. If true, such Imc RFs lead to a computational paradox. On the one hand, vertically elongated Imc RFs are unable to implement stimulus selection in the OT at all possible stimulus location-pairs: selection cannot be solved for over a third of the location-pairs along the elevation (Figure 2.1E). On the other hand, competitive stimulus selection in the OTid is known to occur across all encoded locations [75, 94, 107], with the Imc being the primary source of competitive inhibition [87, 100]. To resolve this paradox, we set out to investigate the functional properties of Imc neurons in the barn owl, as well as the computations implemented by the Imc-OT network in service of stimulus selection across space.

2.1 Multilobed visuospatial RFs in Imc

We measured the visuospatial RFs of Imc neurons using extracellular recordings (section 2.4.4). Individual Imc units were identified by spike sorting single and multiunit data; only those units deemed to be of 'high quality' were included in the analysis (section 2.4.5). Consistent with published data, Imc neurons have high firing rates (median = 76.5 Hz [87, 108]; Figure 2.2AB,EF).

We found that individual Imc neurons possessed visual RFs with multiple, distinct response fields or 'lobes' (Figure 2.2A-H; Figure 2.3AB). The number of lobes in each RF was estimated in an unbiased manner using a two-step process (section 2.4.6): (i) a nonlinear clustering method [109] to fit different numbers of clusters to the spatial map of firing rates followed by (ii) a model selection method [110] to robustly select the optimal number of clusters in the data (Figure 2.2CG, Figure 2.3C-F). We found that about two-thirds of Imc neurons had multilobed RFs (80/116; see also Figure 2.2L).

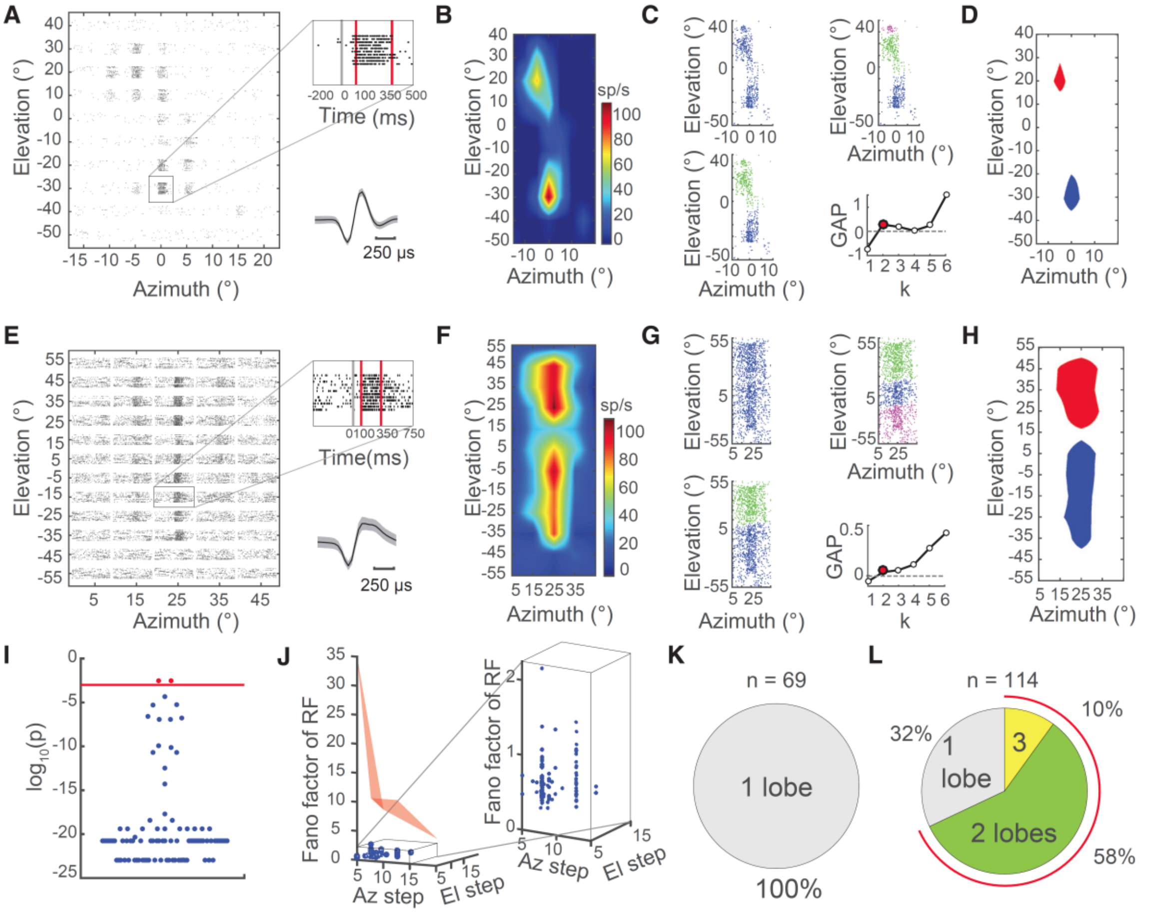


Figure 2.2: Visual receptive fields (RFs) of Imc neurons have multiple distinct response fields ('lobes').
(A) 2-D visual RF of Imc neuron: raster plot of neuron's responses to visual stimulus presented at different spatial locations. ***Inset-top***: Gray line – stimulus onset; red lines – time window used to calculate firing rate; evoked firing rates in Imc were high (median = 76.5 Hz; n=114 neurons). ***Inset-bottom:*** Average spike waveform for neuron in (A); identified as high-quality unit (section 2.4.5); mean (black) ± S.D (gray). (B) Color coded firing rate map corresponding to (A). (C) Rate map in (B) re-plotted as distribution of points in a 2-D plane and subjected to spatial clustering (section 2.4.6). Shown are the best single (*top-left*), best two (*top-right*), and best three clusters (*bottom-left*) fitted to the data using the density peaks clustering method [109] (Figure 2.3C; section 2.4.6). *Bottom-right*: Plot of GAP statistic, a robust model selection metric, against the number of clusters (k) fitted to data [110] (section 2.4.6). Red point: statistically optimal number of clusters (k^*), identified as the smallest k for which GAP exceeds zero; here $k^* = 2$ (section 2.4.6) [110]. (D) Half-max extents of these two optimal RF clusters (lobes). (E-H) Same as A-D, but for a different Imc neuron. (I) Plot of p-values (logarithmic scale) obtained from separability testing for each sorted unit; one-way ANOVA followed by correction for multiple comparisons (section 2.4.5). p-value < 0.05 (blue data): units that are deemed 'well-separated' from co-recorded units as well as outliers (n=114). Red data: units not well separated form cohort. (J) Effect of neuronal response variability and spatial sampling step-size on number of RF lobes detected in a simulated single-lobed RF; Monte-Carlo analysis

(Figure 2.3G; section 2.4.8). Red area: Fano-factor and step-size pairs yielding >5% rate of misidentifying single-lobed RF as multilobed. Blue data: Experimentally recorded Imc neurons (n = 114). (K) Summary of number of RF lobes across 69 OT neurons. See also Figure 2.1 and Figure 2.3. (L) Summary of number of RF lobes across 114 Imc neurons.

To test if the multilobed structure of Imc RFs was an artifact of our experimental or analytical methods, we performed three controls. First, we tested if errors in spike sorting might have caused multiple units with single lobed RFs to be misidentified as a single unit with a multilobed RF. To this end, we applied an additional separability criterion to our sorted units. We tested the statistical separability of the waveforms of each sorted unit with those of any other unit as well as with outlier waveforms recorded at the same site, and retained only those units that were well-separated (section 2.4.5). We found that the majority of the sorted units (114/116) satisfied the separability criterion as well ($p < 0.05$; Figure 2.2I), ruling out multiunit contamination as a source of error. Second, we examined if the spatial sampling resolution used for RF measurement, as well as neuronal response variability, might have caused the erroneous identification of single-lobed RFs as being multilobed (Figure 2.3G). Using experimentally grounded simulations, we mapped out the values of sampling step-size and response Fano-factor that yielded a multilobe misidentification rate of 5% or greater (Figure 2.2J; red zone; section 2.4.8). By comparing with experimental data, we found that the values of these parameters from each recorded unit fell outside the 5% misidentification zone. As a final control, because it is well established that OT RFs have single spatial response fields, we measured visual RFs of OT neurons. Our methods correctly identified all of the measured OT RFs as being single-lobed (Figure 2.2K; Figure 2.3H). Together, these results confirmed the veracity of our conclusion that the Imc contains predominantly ‘multilobe’ neurons (68%; 78/114; Figure 2.2L).

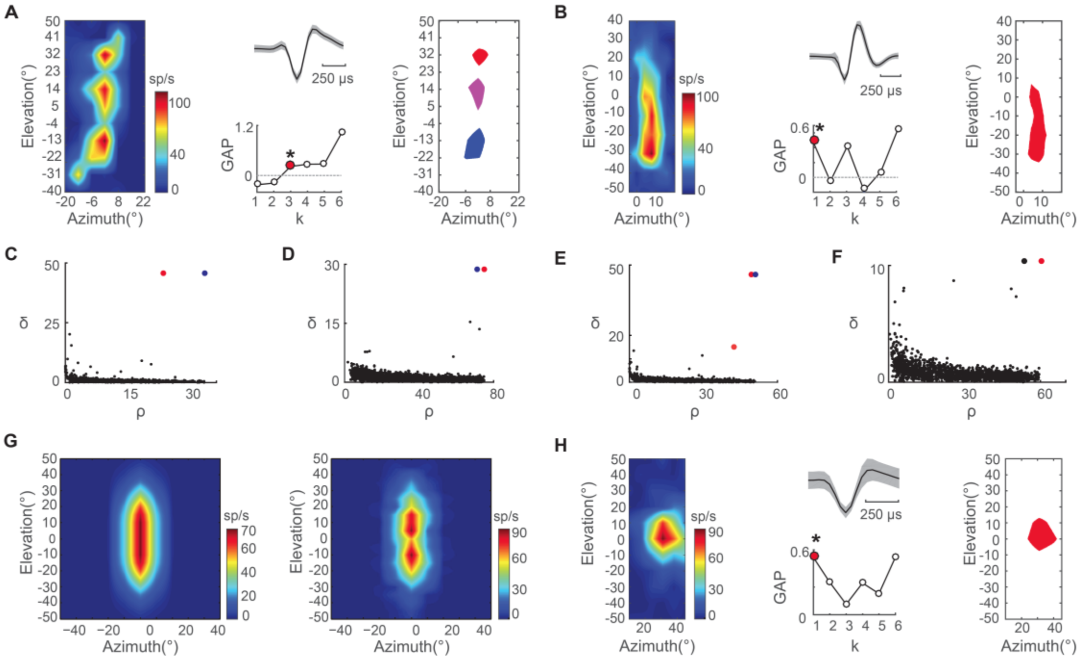


Figure 2.3: Analysis of visual RFs of example Imc and OTid neurons.

(A) Three-lobed visual RF of an example Imc neuron. (Left) Color coded rate map of RF. (Middle, top) spike waveform for the neuron. (Middle, bottom) GAP statistic plot. (Right) Half-max extents of the 3 lobes identified by model selection with the gap statistic. (B) Single-lobed visual RF of an Imc neuron; conventions same as (A). (C) Density peaks clustering method. Scatter plot of local density (ρ) around each data point in Figure 2.2C vs. the distance of that data point from other points that have higher local density (δ). (For the point with highest local density, δ is conventionally taken as the maximum distance of the point from all other points). Points that have both high local density (large ρ value) and that are far away from other points of high local density (large δ value) are potential cluster centers; Red and blue points in this example. Red point corresponds to the center of top cluster, and blue point, the center of lower cluster shown in Figure 2.2D. (D-F) Same as (C), but for RFs in Figure 2.2F, Figure 2.3A and Figure 2.3B respectively. (G) Effect of sampling resolution and neuronal noise on detection of optimal number of lobes in the data. (Left) The simulated single-lobed 2D RF used for the Monte-Carlo analysis in Figure 2.2J (section 2.4.8). Shown are mean firing rates at different locations. (Right) Plot of the RF obtained when it is re-simulated after adding noise (Fano-factor = 30), and sampled with step-sizes = 5° in azimuth and elevation. This sampled RF was identified as having two lobes by our analysis pipeline (conversion to distribution of points on plane, density peak clustering, followed by gap statistic model selection), which is incorrect because the true underlying RF was single-lobed. This illustrates how noisy neural responses may lead to the erroneous conclusion that a single-lobed RF is multilobed. (H) 2D visual RF of an example OTid neuron. Conventions as in (A), (B). The RF is single-lobed.

2.2 Imc encoding along azimuth and elevation

To investigate organizing principles underlying spatial encoding by Imc neurons**,** we analyzed the properties of the measured visual RFs along the two major anatomical axes of the Imc (Figure 2.1A). The azimuthal centers of RF lobes were nearly identical for lobes of individual multilobe neurons (Figure 2.4A, blue data; section 2.4.7), across neurons recorded at a given site (Figure 2.4B, blue data), and across sites recorded along the dorsoventral axis of the Imc (Figure 2.4C; section 2.4.7). However, azimuthal encoding varied systematically along the rostrocaudal axis of the Imc: centers of RF lobes encoded progressively more peripheral azimuths as the recording electrode was moved from rostral to caudal portions of the Imc (Figure 2.4D; [105, 106]).

The encoding of elevation by Imc neurons was strikingly different. RF lobes of individual multilobe neurons were spaced arbitrarily in elevation (Figure 2.4A: large range of red data). Additionally, RF lobes of multilobe Imc neurons were distributed widely across elevational space: for each multilobe neuron (Figure 2.4A, inset: large median of data), across neurons recorded at a given site (Figure 2.4B, red), and across sites recorded along both dorsoventral and rostrocaudal axes (Figure 2.5A-D). There was also no systematic relationship between encoded elevations and distance along either principal axis (Figure 2.5AB).

These results demonstrated that whereas azimuthal space is encoded in a topographic manner along the rostrocaudal extent of the Imc, elevational space is encoded by RFs with multiple, arbitrarily spaced, and widely distributed lobes of varying number and size (Figure 2.5E-J), with a maximum of three RF lobes per neuron (Figure 2.2I).

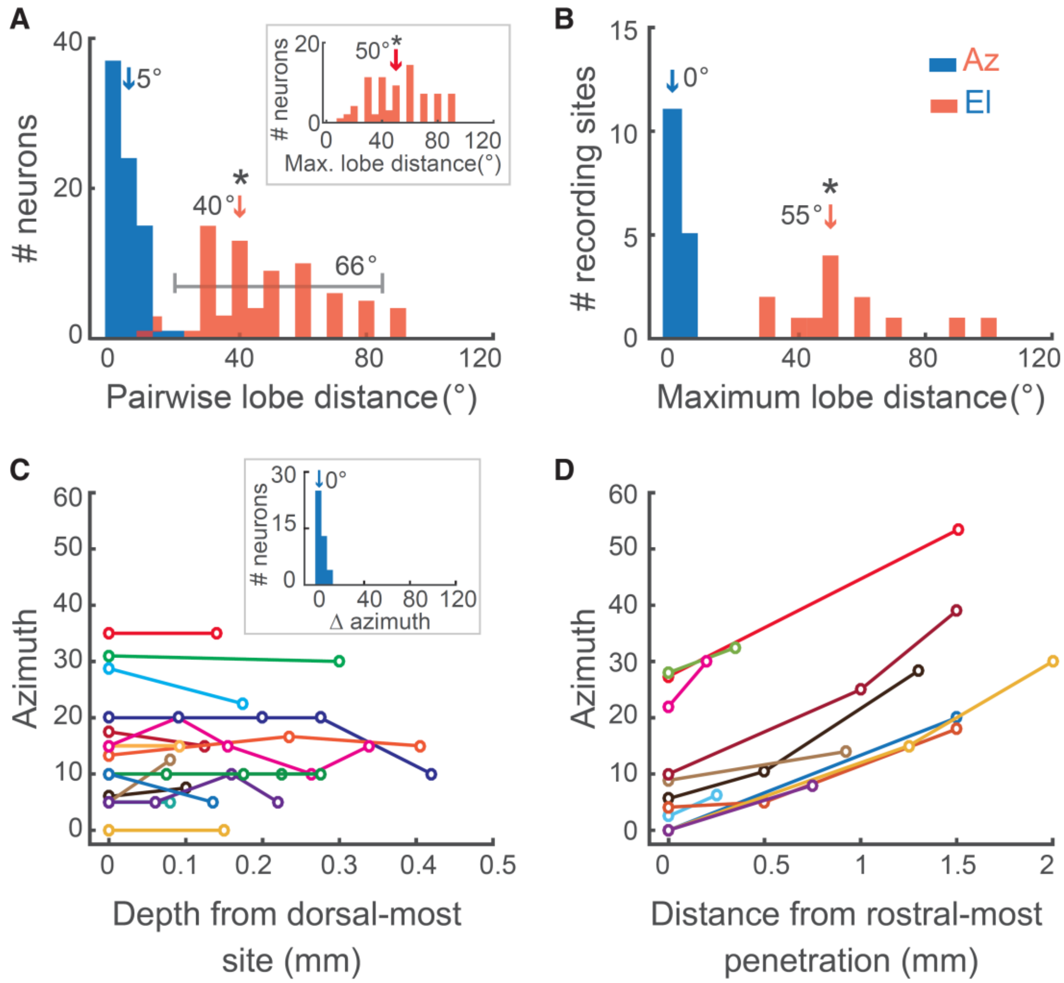


Figure 2.4: RF lobes of multilobe Imc neurons are distributed along elevation but not azimuth, and RFs are organized topographically in azimuth, but not elevation.
(A) Histograms of pairwise distance between centers of RF lobes of individual multilobe neurons (section 2.4.7). Blue: azimuthal distance; red: elevational distance; marked range: 5th to 95th percentile range of red data; large range indicates arbitrary spacing of RF lobes. Arrows: median values; *: median significantly different from 0 ($p = 0.17$, azimuth; $p < 0.05$, elevation; one-tailed ranksum tests). *Inset*: Histogram of maximum elevational distance between centers of RF lobes of an individual multilobe neuron. Arrow: median value; significantly different from 0 ($p < 0.05$, one-tailed ranksum test); large median indicates widely distributed RF lobes. (B) Histograms of maximum distances between centers of RF lobes of multilobe neurons sorted from individual recording sites (section 2.4.5); conventions as in (A); $p = 0.65$ for azimuth; one-tailed ranksum test, $p < 0.05$ for elevation; one tailed t-test. (C) Plot of average azimuthal center of a recording site against the dorsoventral position of the site within the Imc (section 2.4.7); colors: different penetrations. *Inset*: Data re-plotted as histogram of pairwise differences in the azimuthal centers of recording sites along a dorsoventral penetration ($p = 0.18$, one-tailed ranksum test). (D) Plot of average azimuthal 'center' of a dorsoventral penetration against the rostrocaudal position of electrode in the Imc in that recording session (section 2.4.7). Colors: different recording sessions; Spearman correlation =1 in each case. See also Figure 2.5.

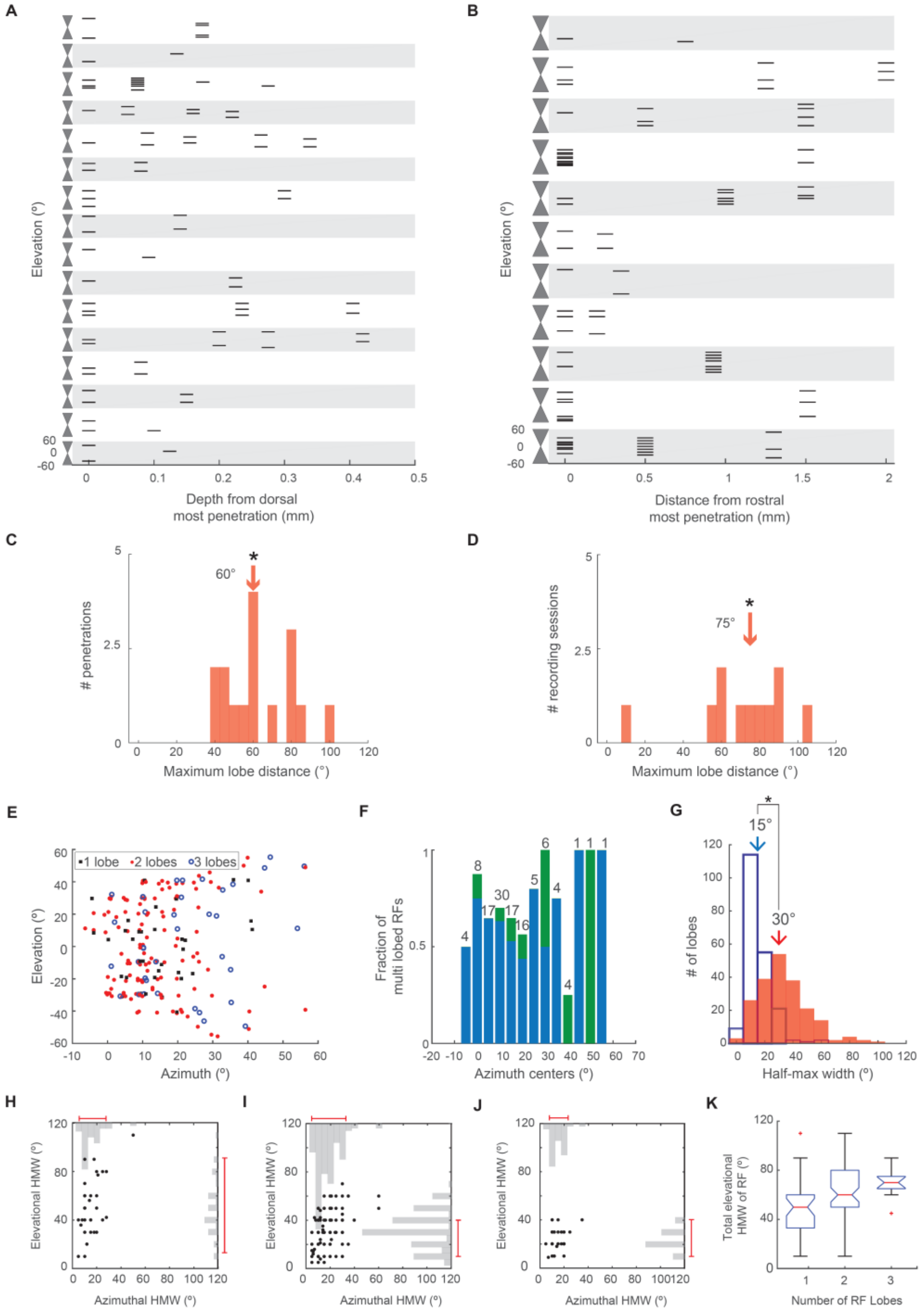


Figure 2.5: Detailed analysis of the organization and structure of RF lobes of Imc neurons.

(A) Plot of elevational centers (black horizontal ticks) of all the visual RF lobes of all individual neurons recorded at a multiunit site, as a function of the dorsoventral position of the electrode within the Imc along a penetration. Each horizontal band (gray and white) band depicts a different penetration; the vertical extent of each band spans -60° to +60° in elevation. No systematic organization of elevational centers of RF lobes along the dorsoventral as evidenced by widespread and irregular distribution of lobe centers at each depth within a penetration. (B) Plot of elevational centers (black horizontal ticks) of all the visual RF lobes of all individual neurons recorded at a multiunit site, as a function of the rostrocaudal position of the electrode during that recording session. Each horizontal band (gray and white) depicts a different recording session; the vertical extent of each band spans -60° to +60° in elevation. No systematic organization of elevational centers of RF lobes along the rostrocaudal axis, as evidenced by widespread and irregular distribution of lobe centers at each penetration (within a recording session). (C) Histogram showing maximum distance between RF lobes measured along each penetration (i.e., each horizontal band in (A)). * indicates mean significantly different from 0 ($p < 0.001$); one tailed t-test. (D) Histogram showing maximum distance between RF lobes measured across all penetrations made along the rostrocaudal axis in each recording session (i.e., each horizontal band in (B)). * indicates mean significantly different from 0 ($p < 0.001$); one tailed t-test. (E, F) Multilobe neurons were found at all tested azimuths. (E) Scatter plot of the azimuthal and elevation centers of the individual lobes of multilobed RFs of all neurons. (F) Fraction of measured RFs that were multilobed, plotted as a function of the azimuthal center of the RF (blue corresponds to 2-lobed Imc RFs and green to 3-lobed Imc RF). (G) Individual lobes of Imc RFs are narrow in azimuth but elongated in elevation. Histogram of azimuthal (open) and elevational (red) half-max widths of all the RF lobes across all recorded neurons. Arrows indicate median values. * indicated that lobes are larger in elevation than in azimuth (ranksum test $p < 10^{-25}$). (H-J) Scatter plot of elevational vs. azimuthal half-max widths of individual lobes of single-lobed RFs (H), two-lobed RFs (I), and three-lobed RFs (J). The azimuthal half-max-widths of the lobes of single-lobed, two-lobed as well as three-lobed RFs are distributed over the same, narrow range of values. (*Top insets in grey*: histograms of azimuthal half-max-widths obtained by projecting data onto the x-axis; ranges of histograms are indicated in red as the extents between the 5th and 95th percentile of the data; HMW: Half-max width.) In contrast, the elevational half-max-widths of individual RF lobes range from the same lower limit to different upper limits for single-, two-, and three-lobed RFs: the tallest single-lobed RFs are taller than the tallest two-lobed RFs, which are in turn taller than the tallest three-lobed RFs. (*Right insets in grey*: histograms of elevational half-max-widths obtained by projecting data onto the y-axis). This result is consistent with the limited extent of encoded elevational space: as the number of lobes in the RF increases, the maximum possible size of the lobes is smaller. (K) Comparison of the total elevational extent of single-, two-, and three-lobed RFs. Total elevational extent of an RF calculated as the sum of the elevational half-max-widths of all lobes of that RF. These data show that total RF height is not a constant across single, two-, and three-lobed RFs. Rather, there are very small as well as very large single- and two-lobed RFs, and three-lobed RFs are larger than most (75% of) single-lobed RFs.

2.3 Neuronal scarcity in coronal planes of Imc

The multilobed encoding of elevational space by Imc neurons was puzzling. This was especially so because neurons that provide input to the Imc (OT10), as well those that receive Imc's output (OTid), all tile sensory space with single-lobed spatial RFs organized topographically in both elevation and azimuth (Figure 2.2K) [77]. Might the implementation of stimulus selection across space, a main function of the Imc [100], impose demands on the spatial coding properties of Imc neurons that can explain multilobed RFs?

To examine the implications of spatial selection on Imc RF structure, and specifically, of the need for implementing stimulus selection at all possible location-pairs, we turned to theory. Briefly, we compared the total number of location-pairs at which selection must occur in the OTid, with the number of location-pairs in the OTid at which selection can be achieved by a set of Imc neurons. Since multilobed Imc encoding is restricted along the elevation (Figure 2.4AB; Figure 2.5A-D), we focused on stimulus selection between all possible pairs of elevations at any azimuth. We proved mathematically that if the number of Imc neurons (N) encoding different elevations at a given azimuth is less than the number of distinct elevational locations (L) encoded by the OTid at that azimuth ($N < L$), then multilobed Imc RFs are necessary for stimulus selection at all possible location-pairs. The details of the mathematical proof that necessitates multilobed encoding in the Imc is described below.

2.3.1 Theoretical calculations regarding the need for multilobed RFs.

We compared the total number of location-pairs at which selection must occur in the OTid, with the number of location-pairs at which selection is achievable by a set of Imc neurons. Since multilobed Imc encoding is restricted along the elevation (Figure 2.4AB), we focused on stimulus selection between all possible pairs of elevations at any azimuth.

2.3.1.1 Simplified version of the proof

We started by making two simplifying assumptions: (a) that the OT space map is a collection of non-overlapping spatial RFs that tile sensory space, and (b) that each Imc neuron has exactly *k* RF lobes (*k* always ≥ 1).

In this scheme, if the number of distinct elevations (at a given azimuth) in the discretized OT space map is *L*, then the total number of distinct pairs of stimulus locations possible is *L(L-1)*. A stimulus placed within any RF lobe of a *k*-lobed Imc neuron can suppress competing stimuli located anywhere outside the RF, i.e., at *L-k* locations. Therefore, each Imc neuron is capable of implementing competitive selection at *k(L-k)* pairs of locations. With *N* such Imc neurons, the number of pairs of stimulus locations at which competitive selection can be resolved by the Imc is at most *Nk(L-k)*. Note that this quantity is computed assuming no overlap between Imc RFs and is greater than the number of pairs of stimulus locations at which competitive selection can be resolved by the Imc if overlap between RFs is allowed. Therefore, to achieve successful competitive suppression between all possible pairs of stimulus locations, i.e., location invariance, a condition that must be satisfied is

$$\mathbf{Nk(L-k) \geq L(L-1)} \qquad \text{- (2.1)}$$

$$\Rightarrow \quad \mathbf{k \geq \frac{L(L-1)}{N(L-k)}} \qquad \text{- (2.2)}$$

This necessary (but not sufficient) condition for location invariance is already very revealing: If all Imc neurons had only single-lobed RFs, i.e., *k* = 1, the above inequality reduces to $N \geq L$ i.e., the number of Imc neurons would need to be greater than or equal to the number of distinct spatial locations. Since the logical proposition 'A => B' is exactly the same as the proposition 'not (B) => not (A)', in our case, the proposition '$k = 1 \Rightarrow N \geq L$' is exactly the same as the proposition '$N < L \Rightarrow k \neq 1$', i.e., if the number of Imc neurons is less than the number of spatial locations, then at least one Imc RF must be multilobed (because RFs cannot have fewer than one lobe, by definition).

This conclusion held true even when both the simplifying assumptions – (a) that OT RFs are non-overlapping, and (b) that all Imc neurons have the same number of RF lobes – were relaxed (see 'Full version of the proof' next).

2.3.1.2 Full version of the proof

We used a more biologically accurate model of space in which RF extents, overlap of RFs across neurons, and the resolution of competition reported in the OTid (the minimum distance between two stimuli such that OTid is able to select the stronger of the two stimuli) [75] were all modeled to match experimental data. In addition, we allowed varying numbers of Imc RF lobes:

Let the total range of elevational locations for which barn owl's midbrain encodes space be R, and the resolution of encoding space be r. Then, the number of distinct locations at which a stimulus can be placed along elevation is $L = \frac{R}{r}$. Let the resolution for competitive selection be C_{res}.

The total number of distinct location-pairs at which two competing stimuli can be placed such that they are greater than C_{res} apart from each other is approximately $L\left(L - \frac{2Cres}{r}\right)$. Note that this quantity is calculated by counting all the locations at which a second stimulus can be placed such that it is at least C_{res} away on either side of a first stimulus that is placed in any of the L locations. However, when a first stimulus is placed at the edge of the visual field, a second competing stimulus can be placed only on one side such that it is C_{res} away. It is straightforward to show that $L\left(L - \frac{2Cres}{r}\right)$ is smaller than the corresponding quantity obtained when edge effects are included. Hence, for location invariance to be achieved, selection of the stronger stimulus must at least be implemented when two competing stimuli are placed in any of these possible location-pairs.

Let the number of lobes in a given Imc neuron be k. Let the half-max size of each lobe be l_h. Then, a k lobed Imc neuron solves competition for a total of $k(L - \frac{l_h}{r}k)$ location-pairs (assuming each Imc neuron sends inhibition to all locations that lie outside the half-max extent of the neuron's RF, without loss of generality; see section 3.6.3 below and Figure 3.1 for implications of this assumption). This is just the number of location-pairs such that one stimulus can be placed inside the multilobed RF (at its peak for effective suppression of competing stimuli) and the other outside. Let the total number of k lobed Imc neurons be N_k.

Therefore, the total number of Imc neurons is

$$\mathbf{N} = \sum_{\mathbf{k}} \mathbf{N_k k} \qquad \text{- (2.3)}$$

To achieve location invariance, we need

$$\sum_k N_k\, k\left(L - \frac{l_h}{r}k\right) \geq L\left(L - \frac{2C_{res}}{r}\right) \qquad \text{- (2.4)}$$

Since $k \geq 1$, and $l_h > 2C_{res}$ (mean l_h = 33.6° ± 1.25° from the 209 RF lobes across 114 Imc neurons we measured, and C_{res} < 10° [75]), we get

$$\left(\boldsymbol{L} - \frac{\boldsymbol{l_h}}{\boldsymbol{r}}\boldsymbol{k}\right) \leq \left(\boldsymbol{L} - \frac{\boldsymbol{2C_{res}}}{\boldsymbol{r}}\right) \qquad \text{- (2.5)}$$

Using (2.5) in (2.4) gives,

$$\sum_{\boldsymbol{k}} \boldsymbol{N_k k} \geq \boldsymbol{L} \qquad \text{- (2.6)}$$

In other words, if all the Imc neurons are single-lobed (k=1), this equation becomes $N \geq L$. Since the logical proposition 'A => B' is exactly the same as the proposition 'not (B) => not (A)', the proposition '$k = 1 \Rightarrow N \geq L$' is exactly the same as the proposition '$N < L \Rightarrow k \neq 1$' i.e., if the number of Imc neurons is less than the number of spatial locations, then at least one Imc RF must be multilobed (because RFs cannot have fewer than one lobe, by definition).

2.3.2 Experimental validation of mathematical proof

2.3.2.1 Along elevation

To examine the biological applicability of this insight, we estimated L and N in the owl brain. For a given azimuth, the OTid encodes elevations ranging typically from -60° to +60°, and does so at a spatial resolution of at least 3° (section 2.4.10). Consequently, the number of distinct elevational locations encoded by the OTid at a given azimuth (L_{el}) is at least 120°/3° = 40 ($L_{el} \geq 40$). Next, we estimated the number of Imc neurons encoding these elevational locations (N_{el}). Because visual azimuth is organized topographically along Imc's rostrocaudal axis (Figure 2.4D), transverse sections of the Imc provide snapshots of Imc tissue encoding all elevations at a given azimuth (Figure 2.6AB). We obtained histological sections perpendicular to the rostrocaudal axis of the Imc and performed Nissl staining to visualize cell bodies (section 2.4.9). Counts of the number of Nissl-stained somata [111] showed that the majority of sections (75%) had fewer than 28 neurons per section (N_{el}; Figure 2.6BC). Thus, N_{el} is typically much smaller than L_{el} (median $N_{el} / L_{el} < 26/40 = 0.65$).

In contrast, along the azimuth, N_{az} is greater than or equal to L_{az}. The OTid encodes azimuths ranging typically from -15° to 60° at a spatial resolution of at best 1° (section 2.4.10). As a result, the number of distinct azimuthal locations encoded by OTid is at most 75 ($L_{az} \leq 75$). On the other hand, we estimated that there are at least 84 neurons involved in encoding these distinct azimuths, $N_{az} \geq 84$ (section 2.4.10). Thus, there are more Imc neurons than there are encoded azimuthal locations ($N_{az} \geq L_{az}$), an observation that is consistent with the absence of multilobed RFs along the azimuth.

These results indicated that multilobed encoding by Imc neurons is consistent with the need for the Imc-OT circuit to achieve stimulus selection at all possible elevational location-pairs in the face of a scarcity of Imc neurons encoding elevation (Figure 2.6BC).

2.3.2.2 Along azimuth

The number of distinct azimuthal locations encoded by OTid was estimated above to be ≤ 75 ($L_{az} \leq 75$).

The rostrocaudal extent of the Imc is 2800 μm, and the somas of Imc neurons are no larger than ~33 μm (largest somatic dimension = 33 μm, n=456 neurons across 20 coronal sections). Therefore, there are at least 84 (coronal) sections along the rostrocaudal axis of the Imc (2800/33), with each section containing at least one Imc neuron not also found in the neighboring sections. (For this conservative estimate of N_{az}, we only need that of the ~26 neurons in each successive coronal section of the Imc (median #neurons per section = 26; Figure 2.6C; dashed red line), just one be distinct.) In other words, there are at least 84 neurons involved in encoding the at most 75 distinct azimuths: $N_{az} \geq 84$. $L_{az} <=75$, which yields that $N_{az} \geq L_{az}$.

Thus, there is sufficient number of Imc neurons to encode azimuthal locations, precluding the need for a combinatorial solution for selection at all location-pairs along the azimuth (involving multilobe neurons with RF lobes spread along the azimuth). Consistent with this expectation, azimuthal encoding by Imc neurons is effectively single-lobed: all lobes of a multilobe Imc neuron encode the same azimuth (Figure 2.6A-C).

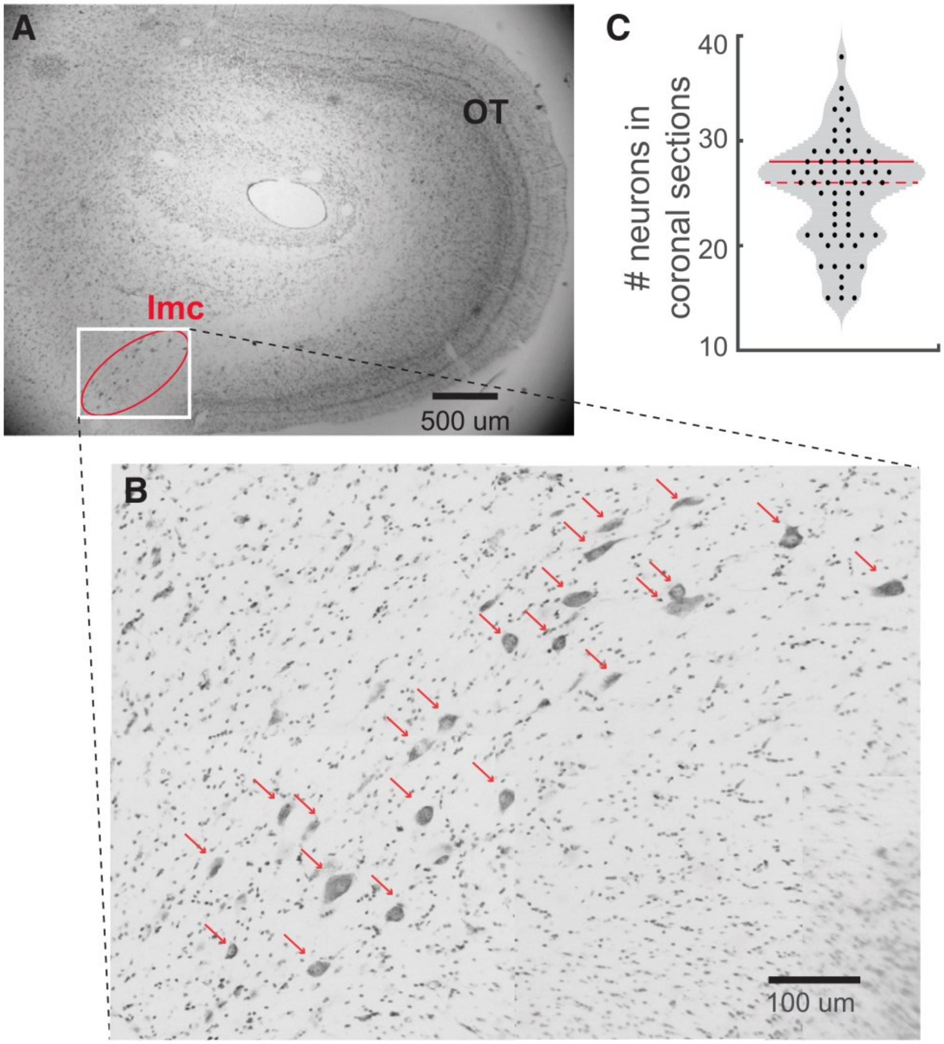


Figure 2.6: Imc encodes elevations with a sparse number of neurons.
(A) Coronal section of owl midbrain showing Imc and OT. **(B)** Zoomed-in image showing individual, Nissl-stained, Imc somata (arrowheads); 22 somata in this section. The zoomed-

in image was obtained by stitching 5 images (each taken at 40x magnification) with overlapping fields of view at the edges. **(C)** Violin plot showing number of Imc somata per coronal section; each dot – one section; n=64 sections across two owls

2.4 Data collection and analysis

2.4.1 Animal care

For the study in chapters 2 and chapter 3, we performed experimental recordings in 15 head-fixed, non-anesthetized adult barn owls that were viewing a visual screen passively (*Tyto alba)*. Both male and female birds were used; the birds were shared across several studies. All procedures for animal care and use were carried out following approval by the Johns Hopkins University Institutional Animal Care and Use Committee, and in accordance with NIH guidelines for the care and use of laboratory animals. Owls were group housed in enclosures within the aviary, each containing up to 6 birds. The light/dark cycle was 12 hrs/12 hrs.

2.4.2 Neurophysiology

Experiments related to chapters 2 and 3 were performed following protocols that have been described previously [75, 100]. Briefly, epoxy-coated, high impedance, tungsten microelectrodes (A-M Systems, 250µm, 5 -10 MΩ at 1 kHz) were used to record single and multi-units extracellularly. A mixture of isoflurane (1.5-2%) and nitrous oxide/oxygen (45:55 by volume) was used at the start of the experiment to anesthetize the bird and secure it in the experimental rig (a 30-minute period of initial set-up). Isoflurane was turned off immediately after the bird was secured and was not turned back on for the remainder of the experiment. Frequently, nitrous oxide was also turned off at this point, but in several experiments, it was left on for a few hours if the bird's temperament necessitated it (some birds were calm when restrained, while others were not). However, it was turned off at least 30 minutes before the recording session. Our recordings were performed starting, typically, 3 hours after initial set-up (the time required for positioning the electrode). As recovery from isofluorane occurs well under 30 minutes after it is turned off, and recovery from nitrous oxide occurs within a minute (the bird stands up and flies away if freed from restraints), recordings were made in animals that were not anesthetized and non-tranquilized.

2.4.3 Imc targeting

The Imc is an oblong structure that is 2.8 mm rostrocaudally and 0.35 mm dorsoventrally, appearing as a 700-µm × 350-µm elliptical disk in coronal sections. It lies parallel to the rostrocaudal axis of the OT, located approximately 16 mm ventral to the surface of the brain, and approximately 500 um medial to the medial-most part of the OT. We targeted the Imc following previously published methods [100]. Briefly we first navigated to the OT (based on well-established methods; [77]), and then navigated to the Imc using the OT's topographic space map as reference. Imc targeting has been validated previously using dye injections [100], and was additionally verified at the outset of this study through anatomical lesions (Figure S1A). Dorsoventral penetrations through the Imc were made at a medial-leading angle of 5° from the vertical to avoid a major blood vessel in the path to the Imc.

2.4.4 Visual stimuli and RF measurement

Visual stimuli used here have been described previously [75, 94]. Briefly, stationary, translating, or looming visual dots (of fixed contrast) were flashed at different locations on a tangent TV monitor in front of the owl. Looming stimuli were dots that expanded linearly in size over time, starting from a size of 0.6° in radius. Visual stimuli were presented for 250ms with an inter stimulus interval of 1.5-3 s at all sampled locations. Pilot experiments indicated that visual RFs were narrow in azimuth but spread along the elevation. Therefore, RF measurements were made by presenting stimuli over the -60˚ to 60˚ range in elevation, and over a 40° (± 10.4°) range in azimuth (centered around the azimuth that yielded the best responses). The ranges of sampling steps used were 4˚-10˚ in azimuth and 5˚-15˚ in elevation. Typically, we sampled 55-70 spatial locations for each neuron (typically, 6 locations along azimuth and 10 locations along elevation). Each sampled stimulus location was repeatedly tested 9-15 times in a randomly interleaved fashion. Multi-unit spike waveforms, recorded using Tucker Davis Technologies hardware interfaced with MATLAB, were sorted off-line into putative single neurons (see below). The spatial responses for each neuron were measured by counting spikes at each sampled location during a 100-350 ms time window following stimulus onset.

2.4.5 Spike sorting multi-unit data

The '*chronux*' spike-sorting toolbox was used for the majority of the analyses [112]. This method is based on a hierarchical unsupervised clustering approach in which the spike waveforms are initially classified into a large number of clusters, typically 10 times the number of putative units recorded. Clusters with very few spikes are discarded and the remaining clusters are then aggregated automatically using metrics of similarity between waveform shapes. In addition, we include only those units for analysis that have less than 5% of the spikes within 1.5 ms of each other (ISI criterion).

The statistical separability of individual sorted units was assessed based on the distance of a unit's cluster (of waveforms) from the clusters corresponding to other units as well as the outlier cluster measured at the same site. We first projected the spike waveforms measured at a given site to a 3-dimensional space using principal components analysis. Then, we performed a one-way ANOVA test to examine if the mean of the waveforms of a given unit (in the projected 3-dimensions) was significantly different from the means corresponding to the other units and the outliers. This was followed by the Holm-Bonferroni criterion for multiple comparisons. In a few cases (4/116), there were either too few waveforms in the outlier cluster (number of waveforms in outlier cluster less than 8% of number of waveforms in any of the remaining sorted units), or the outlier waveforms did not form a cluster with a Gaussian distribution. In such cases, we only tested for the distance of the unit's cluster mean from the cluster means of other units. We regarded only those units whose cluster means were significantly different from the means of all other units (and the outlier cluster) as 'well-separated' units per this separability criterion ($p < 0.05$; the p-value plotted for each unit in Figure 2.2I is the largest p-value obtained across all comparisons for that unit). Only well-separated units were included in all remaining analyses (subsequent to Figure 2.2I) in this study.

2.4.6 Identification of the optimal number of RF lobes

In order to determine the number of response fields (or lobes) in an RF in an unbiased manner, we first transformed the measured RF responses to a distribution of points in 2-dimensional space (azimuth x elevation). This distribution was generated such that the density of points around each sampled spatial location was proportional to the firing rate of the neuron evoked by a visual stimulus presented at that location. We achieved this by distributing points randomly and uniformly within a rectangle centered around the sampled location such that the number of points was equal to the firing rate at that location; the dimensions of the rectangle were the azimuthal and elevational sampling steps, respectively. This transformation allowed us to apply spatial clustering methods to the firing rate maps.

Next, using the density peaks clustering method [109], we fit successively $k = 1,2,3\ldots6$ clusters to the distribution (Figure 2.2CG). This clustering method identifies cluster centers by searching for regions that have high local density of points (ρ) that are also far away from any points of equal or higher density (δ=minimum distance from points of equal or higher density; Figure 2.3C-F. For the point with highest local density, δ is conventionally taken as the maximum distance of the point from all other points). It is robust to nonlinear cluster boundaries and unequal cluster sizes – conditions under which traditional methods like k-means perform poorly. The k cluster centers are chosen by the algorithm as points with the k highest values of gamma (γ), defined as the product of ρ and δ. We repeated this procedure for each k, thereby fitting the 1-best, 2-best, … 6-best clusters to the data.

Following this, we applied a model selection procedure to identify the optimal number of clusters in the data, i.e., the best k value (k*), based on the 'gap statistic' [110]. This is an unbiased method to detect the number of clusters that best fit a distribution of points. For each k, we estimated a 'gap' value (*gap(k)*), which evaluated the goodness of fitting k clusters to the distribution. The gap value was calculated by standardizing the pooled within-cluster sum of square distances between all points in each of the k clusters (W_k) and comparing its log value ($\log (W_k)$) to the expectation of this quantity, ($E^*(\log (W_k))$, under the null hypothesis that the data contains only one cluster [110]. We calculated this in MATLAB by using the *'evalclusters'* function with *'gap'* as the evaluation method, which yielded *gap(k)* as well as *se(k)* for each k; *se(k)* was the standard error in the estimate of *gap(k)*. Then, the gap selection statistic was defined as, $GAP(k) = gap(k) - gap(k+1) + se(k+1)$. The number of clusters that fit the data optimally is defined by the method as the smallest value of k for which $GAP(k) >= 0$. Conceptually, the value of $GAP(k)$ for the null hypothesis ($k^* = 1$) keeps decreasing linearly with increasing k, whereas the rate of the decrease of the metric under the alternate hypothesis ($k^* > 1$) has been shown to fall exactly at $k = k^*$. Hence the *'gap'* between the two curves is maximum at $k = k^*$, and $GAP(k)$, the difference between $gap(k)$ and $gap(k+1)$ is greater than zero for the first time when $k = k^*$.

2.4.7 Defining the centers of RF lobes

The center of an RF lobe defined as the stimulus location evoking the highest firing rate within that lobe. The azimuthal RF 'center' of an Imc neuron is defined as the average of the azimuthal centers of all of its RF lobes, because RF lobes of an individual neuron do not vary significantly in azimuth (Figure 2.4A; blue). The azimuthal RF 'center' of a recording site in the Imc, across all the neurons recorded at that site, is defined as the average of the azimuthal centers across all the RF lobes of all the neurons recorded at that site. This is valid because RF centers of individual neurons within a recording site do not vary significantly in azimuth (Figure 2.4B; blue). The azimuthal RF 'center' of a penetration is defined as the average of the azimuthal centers across all recording sites in that penetration. This is valid because RF centers of individual recording sites within a penetration do not vary significantly in azimuth (Figure 2.4C).

2.4.8 Effect of neuronal noise and spatial sampling on detection of RF lobes

A low spatial sampling resolution during the measurement of spatial RFs, as well as high variability in neural responses, could both cause a single lobed RF to appear falsely as a multilobed one (see Figure 2.3G). To test how robust our method for identifying the ideal number of RF lobes is to sampling resolution (sampling step-size) and neural response variability (response Fano-factor; defined as variance/mean), we performed the following control. First, we generated a single-lobed Gaussian in 2D (azimuth x elevation), with mean and covariance equal to the average values of these parameters across all the experimentally measured Imc RFs (114 Imc units). Using this single-lobed RF as 'reference', we repeatedly simulated RFs using different step-sizes and different response Fano-factor values: For a given step-size, the firing rate at each location was chosen randomly from a normal distribution with mean equal to the value yielded by the reference RF at that location, and variance determined by the Fano-factor value. Next, we transformed this simulated RF into a distribution of 2-D points and applied the density peaks clustering method. Finally, we applied the gap-statistic model selection method to determine the ideal number of lobes in the RF. We repeated this 150 times for each step-size and Fano-factor pair, and calculated the fraction of times for which multiple RF lobes were detected (erroneously) in this data. We repeated the whole procedure for a range of step-size and Fano-factor values that subsumed the range of experimental step-sizes and measured Fano-factor values, and identified the zone that yielded ≥ 5% false detection rate of multiple lobes (Figure 2.2J).

To test the extent to which our experimental and analytical methods falsely detected multilobed RFs, we compared the experimentally used step-size for each RF and the RF's Fano-factor value with those that yielded a ≥ 5% false detection rate in simulation. The Fano-factor for each RF was calculated as the average of the Fano-factor values at all sampled locations in that RF. The step-size for each RF was calculated as the average of the azimuth and elevation sampling steps used to measure the RF. We found that all of our RFs were well within the 'safe' zone of ≤ 5% error (Figure 2.2J). Thus, the detection of multilobed RFs in our data was unlikely to be a spurious consequence of sub-optimal measurement conditions.

2.4.9 Histology

Owls were perfused with paraformaldehyde and their brains extracted per standard procedures. The fixed brains were blocked so that the rostro caudal axis of the Imc was perpendicular to the sectioning plane, and brain sections of 40μm thickness were obtained.

Sections containing Imc were mounted, Nissl stained, and cover-slipped. Sections were imaged at 40x under a light microscope and the number of Nissl stained somata in the Imc in each section were manually counted [111]. For each section, the maximum value of the counts from the two authors was used to generate the plot in Figure 2.6C.

2.4.10 Estimating number of locations encoded by OTid

To estimate, conservatively, the number of distinct elevational locations encoded by the OTid, we divided the extent of elevational space by the poorest resolution of spatial encoding by OTid neurons. The OTid encodes elevations ranging typically from -60° to 60° [77, 100]. The poorest OTid spatial resolution was estimated as the largest distance between adjacent spatial locations for which discriminability (d-prime) of neural responses was above a plausible threshold of 1° [113]. Using published data [45] that reports d-prime values computed from spatial tuning curves in the OTid, we estimated the OTid's poorest spatial resolution at 3°. Thus, the number of distinct elevational locations encoded was at least 120°/3°, in other words, $L_{el} \geq 40$.

To estimate the largest number, the number of distinct azimuthal locations encoded by the OTid, we divided the extent of azimuthal space by the best resolution of spatial encoding by OTid neurons. The OTid encodes azimuths ranging typically from -15° to 60°[75, 77]. In addition, the OTid encodes frontal azimuths with a magnification factor of 150 μm/deg, and peripheral azimuths with a magnification factor of 50 μm/deg [77]. Because resolution varies inversely as the magnification factor, we estimated the best spatial resolution from the poorest using M-scaling $3° * 50/150 = 1°$ [114]. Thus, the number of distinct azimuthal locations encoded was at most 75°/1°, in other words, $L_{az} \leq 75$.

2.5 Discussion

In this chapter, I presented our findings on the visuospatial encoding properties in the Imc. We discovered that Imc uses unusual multilobed visual RFs for encoding elevation space. Such multilobed spatial RFs have not been reported previously in any early visual area to the best of our knowledge. We found that in the Imc, a sensory area that is just two synapses away from the retina [86], the majority of neurons have multilobed visual RFs.

Our report of multilobed RFs in the Imc contrasts with previous reports of large, vertically elongated visual RFs [105, 106]; a consequence of the detailed approaches used here, rather than species differences [87]. Multilobed Imc RFs were characterized here using flashing dots as visual stimuli. The use of this classical approach, which has been employed extensively in visual neuroscience, highlights the contrast between the unusual multilobed encoding of space by Imc and the traditional single-lobed encoding of space by other early visual areas including OT Figure 2.2.

Imc does not play a functional role in OT's representation of the features of single stimuli [100]. In contrast, Imc plays a critical role in delivering global competitive inhibition across the OT space map, and enables the selection of the highest priority stimulus by the OT [87, 100]. Therefore, potential functions for Imc's multilobed RFs were best considered in the context of competitive stimulus selection across space, rather than in the context of neural or behavioral responses to single stimuli. In so doing, we showed that the scarcity of Imc neurons available to encode elevational locations at a given azimuth necessitated multilobed RFs in order to achieve stimulus selection across elevations (in that azimuthal plane).

Given this encoding in the Imc, an important question still remains. How does the scarce Imc implement location-invariant stimulus selection in the OTid space map? Given the asymmetrical encoding of space along the azimuth and elevation, how does stimulus selection operate along these two dimensions? I present our findings that answer these questions in the next chapter.

3. Combinatorial neural inhibition for stimulus selection across space

In this chapter, I present our findings that unravel the neural basis of a critical computation, namely, selection at all possible pairs of locations. This property permits spatial selection to operate no matter which specific locations in the sensory world the competing stimuli occupy. Although appearing straightforward, the implementation of comparisons between all possible pairs of stimulus locations is computationally complex: the number of location-pairs at which two competing stimuli could be placed, $L^2\text{-}L/2$, scales quadratically with L, the number of spatial locations that are encoded. How does the brain meet the resulting demands imposed on neural circuitry and solve stimulus selection at all possible pairs of locations? In the previous chapter, we discovered that the Imc uses multilobed RFs to encode elevational space, that is primarily driven by the scarcity of Imc neurons. In this chapter, using these findings as a basis, we explore how Imc implements location-invariant stimulus selection in the face of such scarcity.

3.1 Optimization model to study stimulus selection in Imc

To explore how an under-complete set of Imc neurons might implement selection at all possible location-pairs, we turned to computational modeling. We set up stimulus selection across spatial locations as an optimization problem with L locations (elevations at a given azimuth), and N model neurons encoding those elevations ($N < L$; Figure 3.1; section 3.2). We imbued all model neurons with Imc-like spatially inverting connectivity with the OT (Figure 2.1 and Figure 3.1). The spatial RFs of these model Imc neurons were represented, for simplicity, using ones and zeros, with ones corresponding to locations inside the RF, and zeros, outside (Figure 3.2B; also see Figure 3.1 for validity of model even when this assumption is relaxed)

The goal of the optimization was to identify the spatial RF structures of these N neurons (i.e., the numbers of their RF lobes and their spatial locations), such that when two stimuli of equal priority were placed at any pair of locations, they suppressed each other equally. This necessary and sufficient condition for implementing selection at all location-pairs was captured by a specially constructed cost function whose value decreased as the number of location-pairs at which the above condition was satisfied increased. The cost function took the lowest possible value of $-L(L-1)$ if and only if the condition was satisfied at all location-pairs (section 3.2). Any set of Imc RFs that achieved this minimum value, i.e., that achieved selection for all location-pairs, was called an 'optimal solution'.

For each value of L, we varied the number of neurons in the model from N = 1 to N = L. In addition, in each case, we examined the impact of single as well as multilobed RFs on the existence and nature of optimal solutions. We did so by including a constraint that specified the maximum number of lobes allowed in a model neuron's RF; denoted by k_{max}. The values of k_{max} tested were 1, 3, and 10, corresponding to key experimentally relevant values: k_{max} =1 only permitted model neurons with (traditional) single-lobed RFs as potential solutions to the optimization problem, k_{max} = 3 permitted up to three lobed RFs, in line with the experimental data (Figure 2.2I), and k_{max} = 10 this allowed up to 10 lobes per RF, representing the largest number of typical Imc RF lobes that one can fit within the encoded elevational space (section 3.6.2). Therefore, the main parameters in the optimization problem were L (number of locations), k_{max} (maximum number of RF lobes allowed per neuron), and N (number of Imc neurons). For each triplet of (L, k_{max}, N), we ran the optimization problem 1000 times (Monte-Carlo simulation), each time with a different, randomly chosen initial condition, to explore the space of potential optimal solutions. The model details are elaborated in the section below.

3.2 Conceptualizing and setting-up the model

3.2.1 Objective function

In our model,

L = number distinct spatial elevations at a given azimuth encoded in our model (i.e., the number of elevations in the 'OTid' space map).

N = number of model Imc-like neurons, i.e., neurons with Imc-like anatomical projection patterns.

k_{max} = maximum number of RF lobes allowed for each model neuron.

*N** = smallest number of upto-k_{max}-lobed model neurons needed to solve selection across *L* elevations.

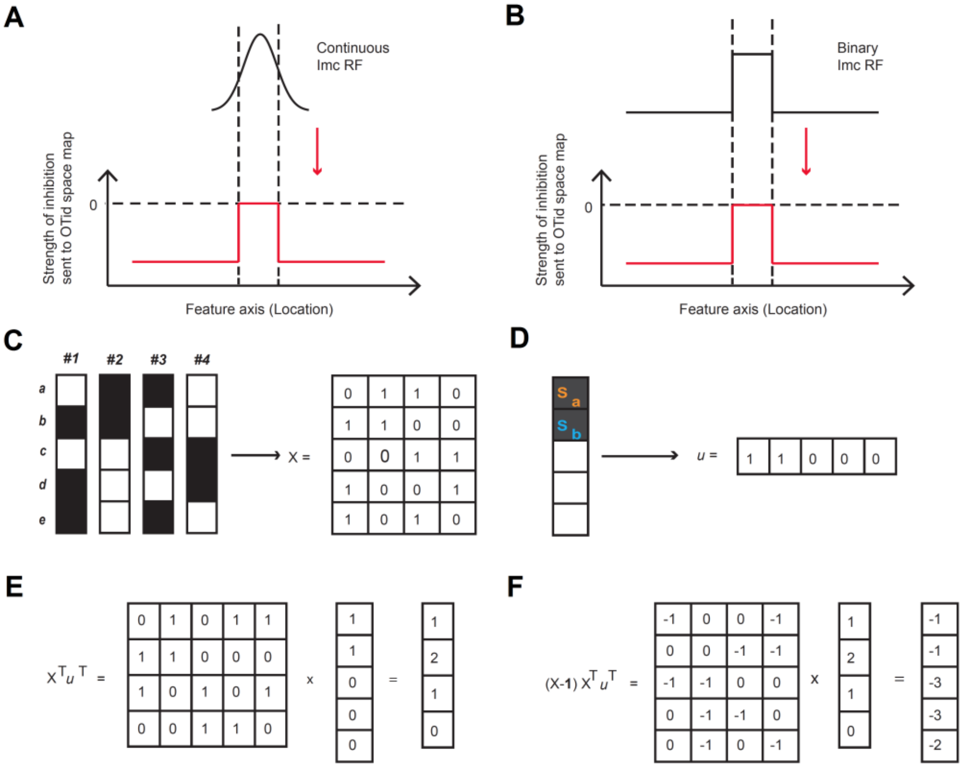


Figure 3.1: Setup of stimulus selection at all location-pairs as an optimization problem.

(A-B) Patterns of spatial inhibition sent by the Imc to OT by continuous Imc RFs (biologically-consistent), vs. binary Imc RFs (simplified abstraction for modeling purposes). (A) Top: Schematic of an Imc RF that encodes locations using continuous values of firing rates. Bottom: Pattern of inhibition sent by the Imc RF to the OTid space map based on the space inverting anatomy between the Imc and OT. Without loss of generality, locations outside the half-max extent of the Imc RF are considered to be spared by Imc spatial inhibition. (B) Same as (A), but when the Imc RF is assumed to be binary at the half-max level of the RF. The spatial pattern of Imc inhibition in (B) is nearly identical to that in (A) (with the exception that the strength of inhibition in (A) gets scaled based on the specific position of the stimulus within the RF half-max.) (C) Left: RF solutions (from Figure 3.2B) obtained by the optimization problem when L = 5 and N = 4. Right: Same RFs represented as an L x N matrix 'X' for the optimization problem. (D) Left: Stimuli presented at locations a and b (from Figure 3.2C). Right: Stimuli pair represented as a row vector u for the optimization problem. (E) The product $X^T u^T$ results in a vector of length N x 1 whose jth element equals the number of locations that activate model neuron j. For instance, neuron #2 is activated by both Sa and Sb. So, the second element of the vector X^Tu^T is 2. (F) The product $(X-1)X^Tu^T$ results in a vector of length Lx1 whose jth element equals the net inhibition sent by the Imc population to location j when the stimuli are presented at locations indicated by vector u. 1 is a matrix, the same size as X with all ones. For instance, the inhibition sent to location b is -1 (by Imc neuron #3). So the second element of $(X-1)X^Tu^T$ is -1.

The optimization model solves for the number and positions of RF lobes of each of the N model neurons in order to achieve selection at all location-pairs. The model neurons are 'Imc-like': each of them is excited by a stimulus placed anywhere within its RF, and delivers competitive inhibition to all locations in the OTid space map outside its RF that is proportional to the strength of the stimulus (Figure 3.1AB). Without loss of generality, we take stimulus priority = 1 unit (for all stimuli), and the proportionality constant (underlying inhibition by the Imc) to be 1. Therefore, for each stimulus, each neuron excited by that stimulus generates an inhibition of 1 unit at those locations in the OTid that are outside that neuron's RF (Figure 3.1AB). For successful, relative-priority dependent competitive stimulus selection between stimuli presented at a given pair of locations, the net inhibition at these two locations in the OTid should be equal. For competitive selection to be solved at all location-pairs, this condition must hold for stimuli placed at any pair of all the possible ${}^{L}C_2$ (L choose 2) pairs of locations. The details of the setup of the optimization problem are described below.

Let X be a matrix of size $L \times N$ (Figure 3.1C), where the j^{th} column of the matrix corresponds to the L elevational locations encoded by the j^{th} Imc neuron in the population. The optimization problem is framed as $\min_X f(X; L, N)$, where the objective function $f(X)$ is designed such that it achieves its minimum value (of $-L(L-1)$) for a given L only when the RFs of the model neurons solve selection at all location-pairs.

Consider two competing stimuli (of equal strength) placed at locations 1 and 2. In our scheme, we represent this by a row vector $u_{1\mathrm{xn}}$= [1 1 0 ….0 …0] (Figure 3.1D). The ones in the first two indices of the row vector correspond to the two locations at which the competing stimuli are placed.

Note that $X^T u^T$ results in a vector in which the j^{th} index corresponds to the number of locations that the j^{th} neuron is activated by when the two competing stimuli are placed in positions shown in a (Figure 3.1E).

Additionally, the matrix (X-**1**) corresponds to the suppression image of the Imc population, where **1** is a matrix vector of all ones. The j^{th} column of this matrix represents the locations to which the j^{th} Imc neuron sends inhibition in the OT space map. This is because of the inverse anatomical projections from the Imc to the OT. The product $(X-\mathbf{1})X^T u^T$ then results in a vector in which the j^{th} index corresponds to the net inhibition sent to the j^{th} location by the entire Imc population when the two competing stimuli are placed at different locations, i.e., at different positions within the row vector u (Figure 3.1F).

For competitive selection at these two locations, the net inhibition at these two locations in the space map of the model 'OTid' should be equal. To penalize solutions for which this is not the case, we include a cost term in the objective function that is equal to the square of difference in the inhibition at the two locations. This is written mathematically as

$$\boldsymbol{f_1(X; u, L, N) = (v(X-1)X^T u^T)^2} \qquad \text{- (3.1)}$$

where v is a row vector whose length equals that of u and nonzero indices are same as u, but with the sign of one of the 1s flipped (in this case v= [1 -1 0 ….0 …0] or [-1 1 0 ….0 …0]). The minimum value that f_1 can take is 0, which happens when equal inhibition is sent to both the locations at which the competing stimuli are placed (Figure 3.1F).

In addition to the strength of inhibition at the two locations being equal, the strength of inhibition must be strictly negative. This is because, the other possibility, of strength of inhibition at each location being zero, would not be acceptable because no inhibition would be sent to either of the two locations. To penalize solutions for which this condition is not met, we include a cost term in the objective function that is equal to the number of locations at which the inhibition is not negative. This is written mathematically as

$$\boldsymbol{f_2(X; u, L, N) = u * sign((X - 1)X^T u^T)} \qquad \text{- (3.2)}$$

Minimizing f_2, therefore, ensures that inhibition is sent to both the locations. The minimum value f_2 can take is -2, when inhibition is sent to both the competing locations (Figure 3.1F).

Finally, we write the full objective for the location-pair (the locations in the pair are specified in the vector u) as below.

$$\begin{aligned} f(X; u, L, N) &= f_1(X; u, L, N) + f_2(X; u, L, N) \\ &= (v(X-1)X^T u^T)^2 + u * sign((X-1)X^T u^T) \qquad \text{- (3.3)} \end{aligned}$$

The minimum possible value that f can take is -2.

For location invariance to be achieved, the function f should be minimized for each pair of locations at which competing stimuli can be placed. In other words, f should be minimized for all possible permutations of vector *u*. This can be written mathematically as

$$f(X; u, L, N) = tr\ (V(X-1)X^T U^T .* V(X-1)X^T U^T) + \ tr\ (U * sign((X-1)X^T U^T))$$

$$\text{- (3.4)}$$

where *U* is the permutation matrix of vector *u* (thus rows of U contain all possible location-pairs). *V* is the corresponding permutation matrix of vector *v*. *tr(Y)* refers to the trace (sum of all the diagonal elements) of the matrix *Y*. Y.*Z is the Hadamard (element-wise) product between the matrices *Y* and *Z* and *sign(Y*) is a matrix obtained by applying the element-wise sign operator to the matrix *Y*.

Because there are ${}^{L}c_2$ possible location-pairs (corresponding to the ${}^{L}c_2$ permutations of the vector a), the minimum value that *f* can achieve is $-2*{}^{L}c_2 = -L(L-1)$. Thus, selection is solved at all location-pairs in our optimization model if and only if the cost function converges to the lowest possible value of $-L(L-1)$.

3.2.2 Constraints

We add two constraints to this optimization scheme. First, we code the RFs of all the model neurons with ones (inside RF) and zeros (outside RF), a simplifying assumption (see section 3.7.1 below for implications of this assumption). Second, we introduce a mechanism to limit the number of lobes in any model neuron to k_{max}. This is done so that, by setting $k_{max} = 3$, we would be able to match the experimentally observed constraint that there are no more than three RF lobes per Imc neuron. The first constraint is fed into the optimization problem as bounded integer constraints with bounds between 0 and 1 to make the RFs binary. The second constraint is implemented as an inequality constraint, written mathematically as

$$\boldsymbol{g(j) = k_{max} - 1_L * X_j \geq 0, for\ all\ j = 1, 2 \ldots N} \qquad \text{- (3.5)}$$

where 1_L is a row vector of length *L*, and X_j is the j^{th} column of *X* corresponding to the RF of the j^{th} neuron. Additionally, we also test the model with $k_{max} = 10$ for some of the analyses reported in Figure 3.2, Figure 3.3, and Figure 3.6.

We solve the above nonlinear optimization problem with mixed constraints, an np-complete problem, using the *'MIDACO'* solver in MATLAB [115].

3.3 Optimal solutions of the model

We found that L = 5 was the smallest number of locations for which selection could be solved at all location-pairs with fewer than 5 neurons (Figure 3.2A). The fewest number of neurons needed by the model in this case, called N*, was 4 (Figure 3.4A; section 3.6.1). Therefore, the maximum 'savings' in the number of Imc-like neurons for L = 5 locations was 1 (= L-N*). We found that as L increased, neuronal savings increased (Figure 3.2A; orange data), with L = 40 locations requiring just N* = 27 neurons to solve selection at all location-pairs (savings of 13 neurons or 32%). Neuronal savings also increased as a function of k_{max}, the maximum number of RF lobes allowed per neuron (Figure 3.2A; black vs. orange data). Notably, when only single-lobed RFs were allowed in the model (k_{max} = 1), N* was always equal to L, and there were no neuronal savings (Figure 3.2A; blue data). Thus, consistent with our theoretical prediction, selection at all possible location-pairs could not be achieved with fewer than L neurons if all neurons only had single-lobed RFs.

The primary motivation for our optimization-based modeling approach was to gain insight into the computational logic underlying successful stimulus selection at all possible location-pairs when neurons are scarce. An example optimal solution obtained when L=5 locations, k_{max} = 3 lobes, and N=4 neurons (Figure 3.2B), illustrates how fewer than L inhibitory neurons can successfully achieve selection at all location-pairs (Figure 3.2CD). Figure 3.2E shows another example optimal model solution, obtained when L = 40 locations, k_{max} = 3 lobes, and N = 27 (N*) neurons.

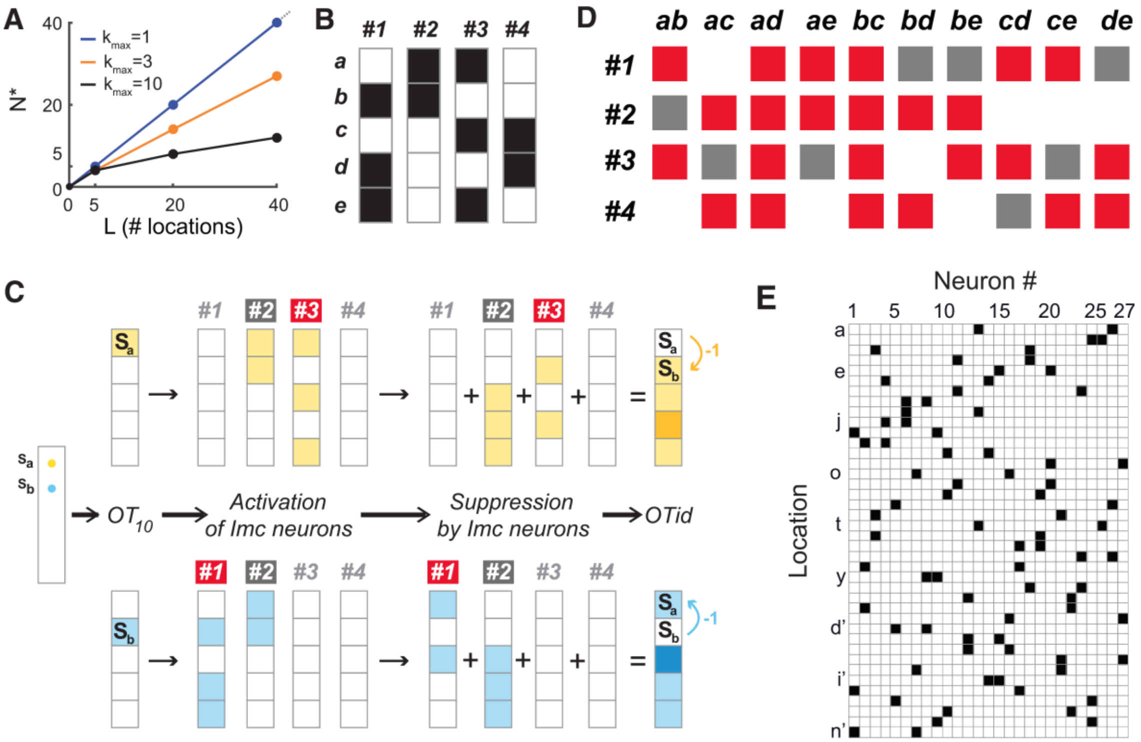


Figure 3.2: Example optimal solution from model illustrates stimulus selection at all location-pairs with an under-complete set of neurons.

Summary plot showing the fewest number of neurons (N*) needed by model to solve selection across all locations for different numbers of locations (L) (see also Figure 3.4A; section 3.6.1). kmax: maximum number of RF lobes allowed for each neuron (section 3.6.2). (B-D) Illustration of selection at all possible location-pairs by an optimal model solution for L=5 locations (numbered a-e) and N=4 neurons (numbered #1-#4). (B) The four RFs in the optimal solution. Shaded areas: RF of neuron; two neurons have multilobed RFs (#1 – two lobes, #3 – 3 lobes). (C) Optimal solution in (B) implements selection between stimuli Sa and Sb at location-pair ab (extreme left). Sa and Sb are of equal priority (1 unit for simplicity). Top row: Information flow through the model OT10-Imc-OTid circuit triggered by Sa. 1st column: Activation of OT10 space map. 2nd column: Activation of individual Imc neurons. 3rd column: Suppression pattern generated by each activated Imc neuron (spatial inverse of the neuron's RF; consistent with published anatomical results; Figure 2.1B-E [85]). 4th column: Combined pattern of suppression in the OTid. Dark colors: 2 units of suppression; light colors: 1 unit (STAR Methods). Curved arrow: Net suppression driven by Sa location b. Dark-gray shading: 'Activated' neuron (#2); defined as a neuron driven by Sa but that does not send inhibition to location b. Red shading: 'Recruited' neuron (#3); defined as activated neuron that sends inhibition to location b, thereby involved in selection for location-pair ab. Bottom row: Same as top row, but for stimulus Sb. (D) Selection matrix summarizing implementation of selection at all location-pairs by optimal model solution in (A). Columns: 10 possible location-pairs; rows: the four neurons. In each column: dark-gray – activated neurons, red – recruited neurons, blank-neurons not activated by either stimulus. (E) Example optimal solution for

(L, kmax, N*) = (40, 3, 27). Black pixels: Locations inside neurons' RF, White pixels: Locations outside neurons' RF.

3.4 Signature properties of optimal model solutions

Detailed analysis of optimal model solutions from *all* runs of *all* (L, k_{max}, N) values tested with N < L revealed that every single optimal solution exhibited three functional properties, which we refer to as signature properties. (Figure 3.3 illustrates these properties for the example optimal solution in Figure 3.2B, and also summarizes them quantitatively for all optimal solutions.)

3.4.1 Multilobe property

First, every optimal solution contained multilobe Imc neurons (Figure 3.2BE, and Figure 3.3A). Conceptually, this 'multilobe property' is necessary because of the scarcity of neurons, i.e., the N < L constraint, as demonstrated previously by theory (section 2.3.1).

3.4.2 Optimized lobe overlap property

Second, the RFs of the neurons in optimal solutions collectively exhibited the 'optimized lobe overlap' property: every multilobe neuron shared each of its lobes, but not all, with another neuron (Figure 3.3B-D). To visualize this property, consider a two lobed neuron M_a (Figure 3.3B: for instance, neuron #1 in Figure 3.2B). There necessarily existed another neuron B (for instance, neuron #2 in Figure 3.2B) in the solution such that the upper lobe of M_a was shared with the lobes of neuron B, but the lower lobe of M_a was not. Similarly, there also existed another neuron C in the optimal model solution such that the lower lobe of M_a was shared with the lobes of neuron C, but not the upper lobe (for instance, neuron #3 in Figure 3.2B). (Here, the neurons B and C could be either single- or multilobed.) This property was quantified using a binary score: briefly, each optimal solution was assigned a score of 1 if every multilobe neuron in that solution satisfied the optimized lobe overlap property (as illustrated in Figure 3.3B), and 0 otherwise. We found that every optimal solution obtained had a score of 1 (Figure 3.3C). Conceptually, the optimized lobe overlap property is necessary because selection needs to be solved also when two stimuli are placed at the locations encoded by different lobes of an individual multilobe neuron (Figure 3.3D). Consequently, this imposes a severe constraint on the relative organization of RF lobes across neurons in optimal solutions - one that causes structured non-orthogonality of the RFs.

3.4.3 Combinatorial inhibition property

Third, neurons in optimal solutions used a 'combinatorial inhibition' strategy to achieve stimulus selection at all location-pairs (Figure 3.3E-J). The combinatorial nature was quantified via a pair of necessary and sufficient conditions, namely that *assorted* subsets of neurons were selectively recruited to solve stimulus selection for individual location-pairs, with the subsets corresponding to different location-pairs *intersecting extensively*.

3.4.3.1 Assortedness property

An optimal solution was said to exhibit the assortedness feature if 'distant' neurons were recruited to solve selection between even nearby locations (Figure 3.3EF), and vice-versa (Figure 3.3EG), *no matter* the specific ordering of neurons in the optimal solution (Figure 3.4B; section 3.6.4.3). This feature was quantified for each optimal solution (section 3.6.4.3): briefly, for a given ordering of neurons in an optimal solution, the selection matrix (as in Figure 3.2D) was constructed, 'nearby' location-pairs were identified (section 3.6.4.3), the largest distance between neurons recruited for selection at each nearby location-pair computed (section 3.6.4.3), and the maximum value of this 'neuronal distance' obtained across all the nearby location pairs in that solution. Then, the minimum value of this neuronal distance was computed across all permutations of neurons within that optimal solution, and across all optimal solutions for that (L, k_{max}, N) triplet (Figure 3.3F; section 3.6.4.3). We found that every optimal solution resulted in the recruitment of distant neurons to solve selection at nearby locations (Figure 3.3F). Conversely, every optimal solution resulted in the recruitment of nearby neurons to solve selection for distant locations (Figure 3.3G; section 3.6.4.3). These results demonstrated the assorted nature of optimal solutions.

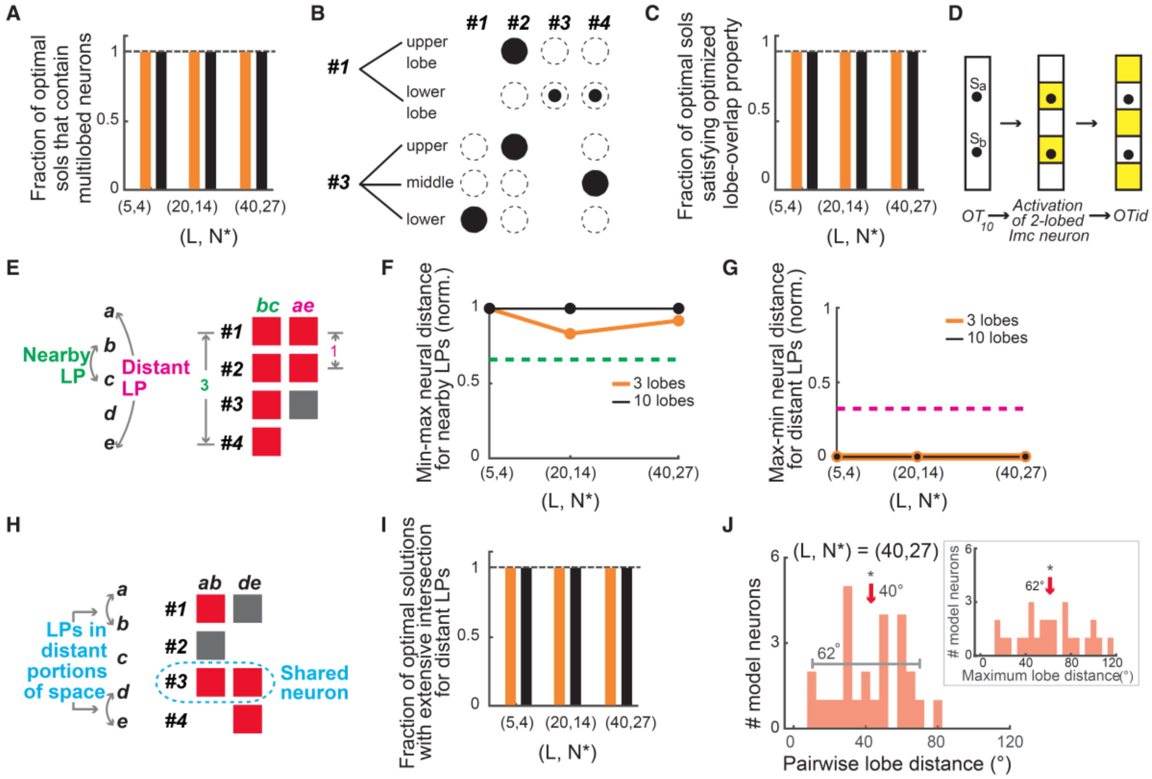


Figure 3.3: Model solutions reveal combinatorially optimized inhibition strategy for stimulus selection at all location-pairs with an under-complete set of neurons. Quantification of signature properties of combinatorially optimized inhibition for optimal model solutions (A) Signature property #1 (multilobe neurons). Fraction of optimal model solutions that had multilobe Imc neurons for all (L, N*) pairs; orange bars – k_{max}=3, black bars – k_{max}=10. (B-D) Signature property #2 (optimized lobe-overlap; see text). (B) Illustration of property for example optimal solution in Figure 3.2B. _Top row_: multilobe neuron *#1* in Figure 3.2B shares upper, but not lower lobe with neuron *#2*, and shares lower, but not upper lobe with neurons *#3* and *#4*. _Bottom row_: Similar, but for multilobe neuron *#3* (see also Figure 3.4E). (C) Fraction of optimal model solutions that satisfy the "optimized lobe-overlap" property; conventions as in (A). (D) Schematic illustrating need for the optimized lobe-overlap property of multilobe Imc neurons. Shown is a two-lobed Imc neuron (*middle*). When stimuli S_a and S_b both occur within the RF of this 'Imc' neuron

(*left*), the resulting zone of suppression generated by this Imc neuron in the OTid would spare both stimuli (*right)*; selection for this location-pair would not be achievable by just this neuron. (E-J) Signature property #3 (combinatorial inhibition; see text). (E-J) 'Assortedness' feature. (E) Illustration of this feature for example optimal solution in Figure 3.2B. *Left panel*: Locations *a-e*. *Right panel*: Patterns of neurons activated and recruited to solve selection for indicated location-pairs (LPs); extracted from selection matrix in Figure 3.2D. Location-pair *bc* involves nearby locations (left panel), but recruits distant neurons to solve selection (right panel; *#1* and *#4*; distance =3, yielding normalized distance of 1, the largest possible value; section 3.6.4.3); conversely, distant location-pair *ae* recruits nearby neurons (*#1* and *#2*; distance =1, yielding normalized distance of 0, the smallest possible value; section 3.6.4.3). (F) Summary plot showing that distant neurons are recruited for selection at nearby locations. Plotted is the normalized distance between neurons recruited for solving selection at nearby locations, termed 'min-max' distance. This is the minimum, taken across optimal solutions and their permutations, of the maximum normalized distance between neurons in a solution recruited for solving selection at nearby locations. Green dashed line: Normalized distance cut-off to be exceeded for neurons to be termed 'distant' (0.66; section 3.6.4.3) and 'nearby' neurons (magenta; 0.33; section 3.6.4.3). (G) Summary plot showing that nearby neurons are recruited for selection at distant locations. Plotted is the normalized distance between neurons recruited for solving selection at distant locations, termed 'max-min' distance. This is the maximum, taken across optimal solutions and their permutations, of the minimum normalized distance between neurons in a solution recruited for solving selection at distant locations. Magenta dashed line: Normalized distance cut-off to not be exceeded for neurons to be termed 'nearby' (0.33; section 3.6.4.3). (H-J) 'Extensive intersection' feature. (H) Illustration of this feature for example optimal solution in Figure 3.2B: location-pairs occupying distant portions of space (left panel) recruit intersecting neural subsets to solve selection (right panel; section 3.6.4.3). See also Figure 3.4. (I) Fraction of optimal model solutions that satisfy the extensive intersection feature (section 3.6.4.3; conventions as in (A). (J) Lobes of neurons in model solutions were arbitrarily placed and widely spread. Histogram of pairwise distance between centers of RF lobes of individual multilobe neurons for a randomly selected model solution for (L, k_{max}, N*) = (40, 3, 27); marked range: 5th to 95th percentile range of red data; large range indicates arbitrary spacing of RF lobes. *Inset*: Histogram of maximum distance between centers of RF lobes of a multilobe neuron, for the same model solution. Arrow: median value; significantly different from 0 ($p < 0.001$, one-tailed ranksum test); large median indicates widely distributed RF lobes.

3.4.3.2 Extensive intersection feature

An optimal solution was said to exhibit the extensive intersection feature if neural subsets recruited to solve selection even for two location-pairs in distant portions of space shared common neurons (Figure 3.3HI). This feature was quantified by first identifying 'doublets':

two location-pairs such that the locations within each pair were nearby locations, but such that the two pairs themselves occupied distant portions of space. Then we checked if the neural subsets recruited to solve selection for at least one such doublet involved a common neuron, and scored the solution as 1 if they did (section 3.6.4.3). We found that optimal solutions obtained from all runs exhibited this feature (Figure 3.3I), demonstrating the extensively intersecting nature of optimal solutions. Together, the above results indicated that combinatorial inhibition was a signature property of optimal solutions.

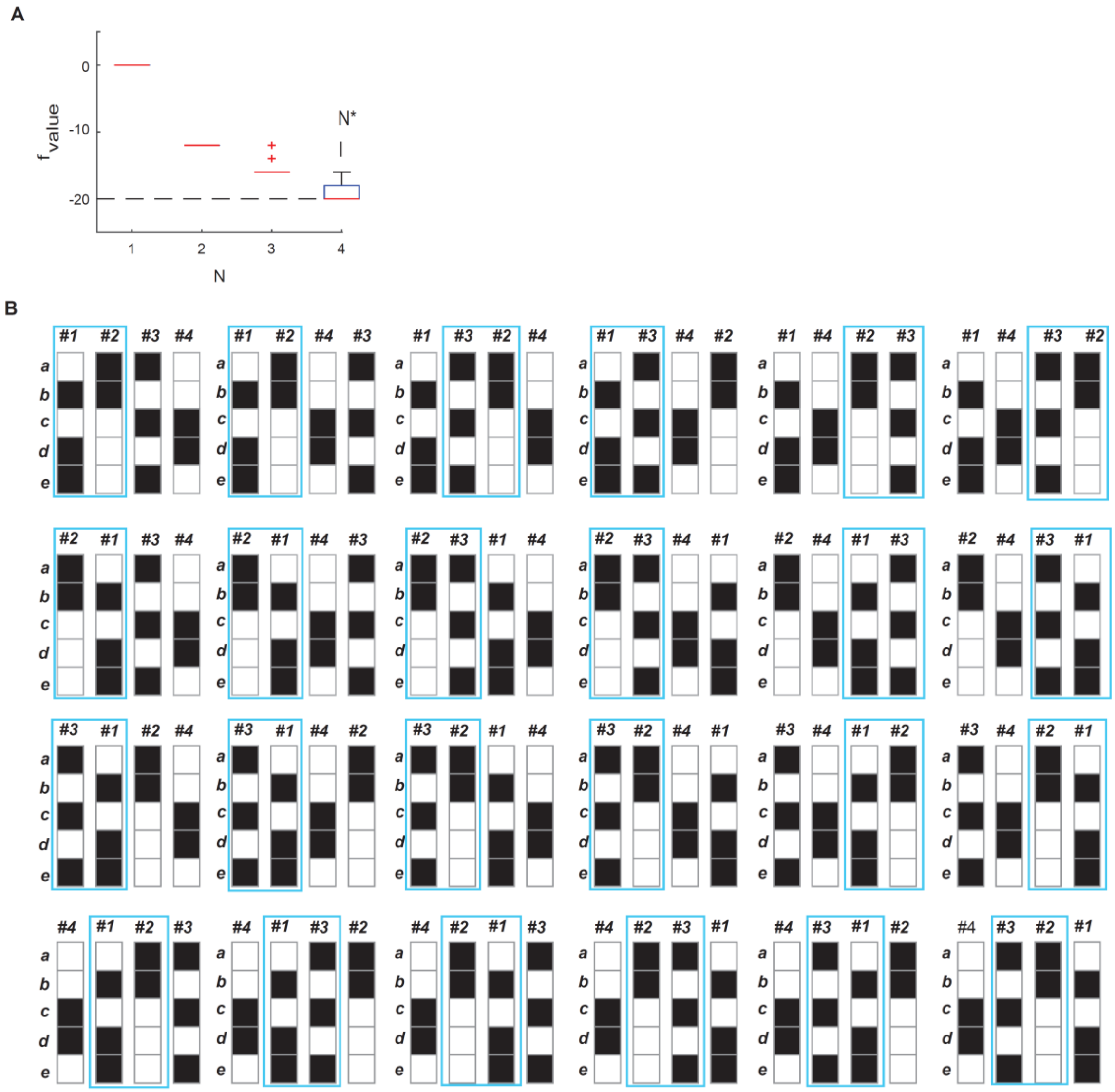


Figure 3.4: Optimal model solutions and combinatorial coding of space.

(A) Minimum value of cost function achieved by the optimization model with L = 5 locations, plotted as a function of number of Imc-like neurons in the model; optimization was run 1000 times for each N. The minimum value progressively decreased as N increased, achieving the lowest possible value that the cost function can achieve (-L(L-1); -20 for L = 5) only when N = 4. In other words, the smallest number of neurons at which selection can be achieved by the model at all possible location-pairs, called N*, was 4 when L = 5 locations. Therefore, neuronal savings, defined as L-N*, was 1. (B) All 24 possible permutations of the model solution in Figure 3.2B; same conventions as in Figure 3.2B). For each permutation, there is at least one pair of nearby neurons that encode distant locations (indicated by blue box).

Taken together, the model revealed that selection at all possible location-pairs when N < L, as is the case with Imc's elevational coding, necessitates a combinatorially optimized inhibition strategy by multilobe neurons.

In contrast, when N ≥ L, as is the case with Imc's azimuthal encoding, the model was always able to solve selection at all location-pairs with just single-lobed neurons (Figure 3.2A, k_{max}=1, blue data), by using the straightforward 'modular copy-and-paste' strategy (Figure 2.1D).

3.5 Experimental validation of model predictions in Imc

The above modeling results indicated that sparse sets of Imc-like inhibitory neurons use a combinatorially optimized inhibition strategy to achieve selection at all possible location-pairs. However, it is unclear if the owl Imc does, in fact, implements this strategy for selection across elevational locations. To examine this, we tested whether the experimentally recorded activity of Imc neurons exhibited the three signature properties of combinatorially optimized inhibition predicted by the model. Because all elevations at a given azimuth are encoded by neurons within a coronal plane (Figure 2.4BC), we sampled these neurons by making recordings at multiple dorsoventral sites within each coronal plane (section 3.6.4).

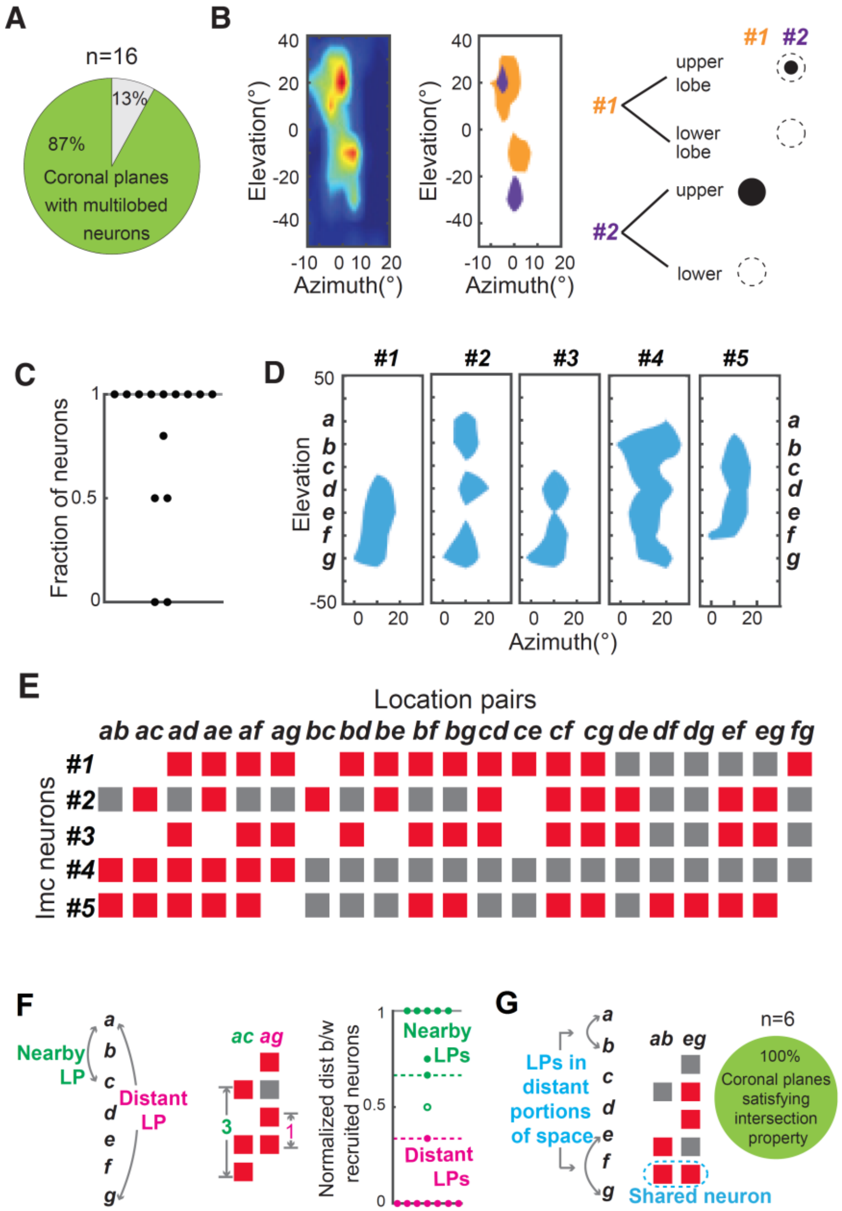


Figure 3.5: Experimental validation of combinatorially optimized inhibition in the Imc.
(A) Signature property #1: Pie-chart summary of fraction of Imc coronal planes tested that contained multilobe neurons (87% = 14/16 planes). (B-C) Signature property #2. (B) Left: Rate map of RF of another Imc neuron sorted from the same recording site as the neuron in Figure 2.2A-D. (Only these two neurons were recorded in this Imc coronal plane.) Middle: Half-max of RFs of neurons in Figure 2.2 (purple; reproduced from Figure 2.2D) and Figure 3.5B left (orange). Right: For each neuron, the upper RF lobe, but not lower one, shows overlap, satisfying the testable lobe-overlap property (see text); conventions as in Figure 3.3B. (C) Fraction of multilobe neurons in each coronal plane satisfying the testable version of lobe-overlap property; dot – coronal plane; median fraction = 1. (D-F)

Signature property #3. (D) RFs (half-max) of all Imc neurons recorded within an example coronal plane. a-g are seven (discretized) spatial locations encoded by these neurons. (E) Selection matrix showing combinatorial activation of recorded neurons for selection at different location-pairs; conventions as in Figure 3.2C. (F) Two left panels: Illustration of assortedness feature for example in (D); conventions as in Figure 3.3E (section 3.6.4.3). Right: Summary of this feature across Imc coronal planes; only those planes containing ≥ 3 Imc neurons each were testable (8/14; section 3.6.4.3) Dashed lines: Distance cut-offs for 'distant' neurons (green; 0.66) and 'nearby' neurons (magenta; 0.33; section 3.6.4.3). Filled circles: Imc coronal planes that satisfied these cut-off criteria; ≥ 7/8 in each case (STAR Methods). (G) Left: Illustration of 'extensive intersection' feature for example in (D); conventions as in Figure 3.3H. Right: Pie-chart summary of this feature across coronal planes (100% exhibited the feature; 6/6). Note that this feature was testable only for those planes for which the recorded neurons encoded location-pairs occupying distant portions of space (6/14; section 3.6.4.3).

Across recordings made in 16 such coronal planes, we found that multilobe neurons were present in nearly every case (14/16; Figure 3.5A; also Figure 2.5A), thereby validating signature property #1. The impracticability of recording exhaustively from all Imc neurons in a coronal plane made it infeasible to test if *every lobe* of each multilobe neuron satisfied the optimized lobe-overlap property (signature property #2; Figure 3.3B-D). Therefore, we tested if *at least one lobe* of each multilobe neuron satisfied it (Figure 3.5B; section 3.6.4.2). The median fraction of multilobe neurons in each coronal plane that satisfied signature property #2 was 1 (Figure 3.5C).

Finally, we tested for signature property #3 (combinatorial inhibition). Both its features, namely, assorted recruitment and extensive intersection, were satisfied in nearly every testable case (7/8 and 6/6 planes respectively; Figure 3.5D-G; section 3.6.4.3), despite the non-exhaustive sampling of Imc neurons in individual planes. In addition, the computational basis of combinatorial inhibition, namely, combinatorial coding of space by model Imc neurons, which was quantified via the arbitrary spacing and wide distribution of their RF lobes (Figure 3.3J), was also found in experimental data (Figure 2.4AB, Figure 2.5A-D).

Thus, our experimental results show that owl Imc neurons are activated in a combinatorially optimized fashion that solves stimulus-priority dependent competitive selection across locations-pairs in the OT.

3.6 Analysis methods for model (3.4) and experimental validation (3.5)

3.6.1 Estimating *N**

*N** is the smallest number of model neurons needed to solve selection at all location-pairs for a given *L* and k_{max}, i.e., the smallest *N* for which the minimum value of the objective function (*-L(L-1)*) can be successfully achieved. This was estimated as follows. For each value of *N* from 1 to *L*, we ran the optimization model 1000 times (1000 runs). Any given run was said to have converged to a solution if the value of cost function did not change for 1000 successive iterations (by setting the *'evalstop'* criterion in the optimization code to 1000), thereby reaching an asymptotic value. The collection of model neuron RFs at convergence was called a 'convergent solution'. Additionally, if the convergent solution attained the value of *–L(L-1)*, then it was called an 'optimal solution'. In other words, optimal solutions are ones that converged and additionally achieve stimulus selection at all location-pairs.

*N** (for a given *L* and k_{max}) is, therefore, the smallest value of *N* for which at least one of the 1000 runs yielded an optimal solution, meaning that for $N = N^*$-1, none of the 1000 runs yielded a solution that successfully achieved selection across all L locations.

For instance, if k_{max}=1 lobe, then for all *L*, $N^* = L$ (Figure 3.2A, blue data; consistent with theoretical calculation presented in the 2.3.1). If k_{max} = 3 lobes and *L* = 5 elevations, all runs for all values of N from 1 to *L* yielded convergent solutions, but *optimal* solutions were produced only when $N \geq 4$ (Figure 3.4A). More generally, if $k_{max} > 1$ lobe, then for all $L > 4$, $N^* < L$ (Figure 3.2A; orange and black data).

3.6.2 Range of k_{max} values chosen for analyses

The specific values of k_{max} used in our simulations (Figure 3.2, Figure 3.3 and Figure 3.6) were 1, 3, and 10 lobes. The reasoning for this choice of values is described below.

$k_{max} = 1$ lobe corresponded to the null hypothesis of single-lobed RFs

$k_{max} = 3$ lobes represented Imc data (Figure 2.2L)

$k_{max} = 10$ lobes. (i) The range of elevations encoded by the OTid and the Imc is no greater than -60° to -60°, and (ii) Most individual RF-lobes have a half-max height ≥ 10° (10-percentile value of half-max height of an individual RF lobe = 10° (Figure 2.5G). Therefore, the number of possible distinct lobes along elevation for RFs of typical Imc neurons ≤ ~10 lobes (=120°/(10° + 2°); with the two added degrees representing 1° spacing on either side of a lobe to separate it from abutting ones.)

3.6.3 Model assumptions

Our optimization model makes two key simplifying assumptions: (a) discretized (pixelated) spatial locations, and (b) binary (on or off) RFs of the model neurons. The former assumption can be readily reconciled with biology by making the pixel size sufficiently small. Therefore, this assumption does not result in loss of generality of the model. Second, the pattern of spatial inhibition sent to the OTid space map, the key computational function required of Imc in the model, is the spatial inverse of the RF: inhibition is sent to all locations except the ones inside the RF. In other words, the spatial pattern of inhibition is, by definition, a '*binarized* spatial inverse' of the Imc RF, with the strength of delivered inhibition being proportional to the specific location within the continuous RF at which the stimulus is placed (Figure 3.1AB). For the model, it is the pattern of inhibition that is critical, informationally speaking, rather than the variations in the strength of delivered inhibition based on the specific location within RF that a stimulus occupies (Figure 3.1AB). (This is unlike population vector coding, where the specific values of firing rates within an RF are critical informationally [116-119]). Therefore, the continuous RF can be binarized itself (say, at the half-max, or 75%-max level) without the qualitative conclusions of the model being affected (Figure 3.1AB). Notably, despite these simplifying abstractions of the biology by the model, we found that predictions from the model held true experimentally (Figure 3.5), further revealing that the model captured sufficiently well the key computational principles at play in this circuit. Consequently, it was able to provide a compelling explanation for the unusual functional properties of Imc neurons, and illuminate the neural mechanisms by which this midbrain circuit solves stimulus selection at all location-pairs.

3.6.4 Characterizing signature properties

3.6.4.1 "Multilobe property" (property #1).

Model: For each optimal solution at each (L, k_{max}, N) tested, we examined if any of the model RFs were multilobed. A model RF was said to be multilobed if it had "on" pixels that were separated by "off" pixels; two adjacent "on" pixels were treated as one lobe. For instance, in Figure 3.2B, neurons #2 and #4 have one lobe each. Neuron #1 has two RF lobes and neuron #2 has 3 RF lobes. These two neurons are multilobed. Thus, this optimal model solution is said to satisfy the "multilobe property".

Data: For each coronal Imc plane recorded, we examined if any of the neurons in that plane had multilobed RFs. Whether an RF was single or multilobed was determined using methods described in (and surrounding) Figure 2.2.

3.6.4.2 "Optimized lobe-overlap property" (property #2).

Model: A multilobed model neuron that shares each of its RF lobes, but not all, with another neuron is said to satisfy this property. If every neuron in a model solution satisfies this property, the model solution itself is said to satisfy the optimized lobe-overlap property. The fraction of model solutions satisfying this property for each (L, N*) is plotted in Figure 3.3C (100%, in each case).

Data: The set of neurons recorded within a given coronal plane, i.e., across all the recording sites along a dorsoventral penetration, is collectively a potential solution set for solving selection across all elevation pairs at that azimuth. (This is because of our finding that spatial azimuth is encoded topographically along the rostrocaudal axis of the Imc, and all the elevations at a given azimuth are encoded by the neurons in the coronal plane at the appropriate point along the rostrocaudal axis; Figure 2.4 and Figure 2.5). A multilobe neuron that shares at least one of its RF lobes, but not all, with another neuron in the solution set is said to satisfy the experimentally testable version of the lobe-overlap property. To test this property in data, we first obtained the set of discrete elevational locations encoded by Imc neurons in a solution set (coronal plane). We did this by quantizing, at a resolution of 3° (to match theory and model; see main text related to Figure 2.6), the maximum elevation range encoded by their RFs combined. Next, an RF lobe of a multilobed Imc neuron was said to overlap with the RF of another neuron if there existed a location within the former's half-max extent that also lay within the half-max extent of the latter's RF. The fraction of multilobed

Imc RFs in each coronal plane that satisfy this testable version of the optimized lobe-overlap property is shown in Figure 3.5C. (This testable version of the lobe-overlap property was necessary because of the inherent infeasibility of recording from all Imc neurons in a coronal section, i.e., from all the neurons in a 'solution set'. Specifically, the small mediolateral extent of the Imc (< 350 μm), coupled with the thickness of the electrode (250 μm) that was used to reliably target the deep Imc (~16 mm below brain surface), limited us to one dorsoventral penetration within a coronal section. This made recording from all Imc neurons in a given section unviable. The average # neurons recorded per section = 3.44 ± 0.47.

3.6.4.3 'Combinatorial' property (property #3).

(A) ***<u>"Assorted neural subset" feature</u>***: Distant neurons are recruited to achieve selection for nearby locations, and nearby neurons are recruited to achieve selection for distant locations. To test for this feature, we divide the elevation range (L locations) into three parts, the upper L/3, middle L/3 and lower L/3 locations. Two locations are said to be 'nearby' if the distance between them is ≤ L/3, and 'distant' if the distance between them is ≥ 2L/3. Similarly, two neurons are said to be nearby if the distance between them is ≤ (N-1)/3, and distant, if their distance is ≥ 2(N-1)/3. We then ask if distant neurons are recruited for a nearby location-pair (LP), and vice-versa. Since there is no meaningful functional ordering of multilobe neurons owing to the lack of topography in the encoding of elevation, we must test these questions across permutations of the ordering of Imc neurons within a solution.

i. ***<u>Model</u>***: First, we tested if distant neurons are recruited for a nearby location-pair. We did so by computing the following metric (Eq. 3.6) for each (L, N*) (Figure 3.3EF).

$$\boldsymbol{d\ (nearby\ LP) = \left[\min_{solutions}\left\{\min_{permutations}\left(\max_{nearby\ LP}(d)\right)\right\}\right]} \qquad \text{- (3.6)}$$

Here, ‘d’ is the maximum distance between the neurons recruited for solving selection for a given nearby location-pair in a given solution. The maximum of this across all nearby location-pairs yields the farthest distance between neurons recruited to solve selection for any nearby location-pair. The minimum of this value across permutations of neurons in the solution, and across all solutions, yields d (nearby LP) for that (L, N*).

For L = 5 (N* = 4), we tested this exhaustively for all possible permutations (= factorial(4)). However, for L = 20 (N* = 14) and L = 40 (N* = 27), the number of permutations is very large ($14! = 8.7 \times 10^{10}$ and $27! = 1.08 \times 10^{28}$). Because it was infeasible to test all possible permutations in these cases, we tested a subset of permutations (n=1000) that was selected randomly from the set of all the possible permutations using the ‘randperm’ function in MATLAB.

For each (L, N*), we calculated the normalized minimum distance between neurons recruited for selection at distant location-pairs as shown in (Eq. 3.7), and plotted it in Figure 3.3F.

$$\boldsymbol{d_{norm}(nearby\ LP) = \frac{d(nearby\ LP) - d_{min}}{d_{max} - d_{min}}} \qquad \text{- (3.7)}$$

Here, d_{max} (= N*-1) and d_{min} (= 1) are the maximum and minimum possible distances between neurons in a solutions set consisting of N* neurons. We found that in every case, this normalized distance was high (> 0.66; the normalized cut-off value chosen for defining ‘distant’ neurons).

Next, we tested if nearby neurons are recruited for a distant location-pair (Figure 3.3EG), using a metric constructed with a logic similar to that used above:

$$\boldsymbol{d\ (distant\ LP) = \left[\max_{solutions}\left\{\max_{permutations}\left(\min_{distant\ LP}(d)\right)\right\}\right]} \qquad \text{- (3.8)}$$

$$\boldsymbol{d_{norm}(distant\ LP) = \frac{d(distant\ LP) - d_{min}}{d_{max} - d_{min}}} \qquad \text{- (3.9)}$$

For each (L, N*), we calculated the normalized maximum distance between neurons recruited for selection at distant LPs (Eq. 3.9), and plotted the results in Figure 3.3G. We found that in every case, this normalized distance was small (< 0.33; the normalized cut-off value chosen for defining 'nearby' neurons).

ii. ***Data***: For Imc neurons in each solution set (coronal plane), we obtained the range of discretized elevation values encoded as before (resolution of 3°), and then calculated the normalized minimum distance between nearby neurons and the normalized maximum distance between distant neurons using the Eqns. (3.7) and (3.9) above. Note that for the notions of nearby neurons and distant neurons, there need to be at least 3 neurons in the solution set so that the maximum distance is 2 and the minimum distance is 1. Out of 14 coronal planes that contained multilobe neurons, 8 had ≥ 3 neurons. The results from these 8 planes are plotted in Figure 3.5F.

(B) ***"Extensive intersection" feature***. Location-pairs occupying distant portions of space recruit shared neurons to solve selection at each pair. Two location-pairs are said occupy distant portions of (elevational) space if one location-pair lies within the upper third of the locations (upper L/3) and the other lies within the lower third of the locations (lower L/3). Since intersection between the neural subsets is independent of the ordering of the neurons, we do not need to test this for all permutations of neuron orderings.

i. ***Model***: For every optimal solution at a given (L, N*), we tested if there existed two location-pairs (a 'doublet') occupying distant portions of space such that the neural subsets recruited to solve selection at each location-pair shared at least one neuron. The fraction of optimal solutions that satisfied this property is plotted as a function of (L, N*) in Figure 3.3I; the fraction is uniformly 100%.

ii. ***Data***: For Imc neurons in each solution set (coronal plane), we obtained the range of discretized elevation values encoded as before (resolution of 3°). We then tested if these neurons satisfied the extensive-intersection property as described for the model. Of the 14 coronal planes at which neurons were recorded, in 6 cases, the encoded locations included two location-pairs that occupied distant locations. The fraction of these 6 coronal planes that satisfied the extensive intersection property is shown in Figure 3.5G (100%).

3.7 Discussion

This study unpacked the mechanistic underpinnings in owls of a critical neural function, namely stimulus selection at all location-pairs, and revealed combinatorially optimized inhibition as an efficient strategy for it when neurons are scarce. Combinatorial activation of neurons is typically inferred by visualizing their patterns of activation in response to stimuli; for instance, combinatorial activation of odors by olfactory receptor neurons; [120]. Here, in addition to visualizing these patterns, we quantified the underlying intuition by defining the 'assortedness' and 'extensive intersection' features, which are necessary and sufficient for combinatorial activation, and applied this approach to both model solutions as well as experimental data. This allowed for quantitative support for the finding of combinatorially optimized activation of inhibitory neurons. Broadly, this study was framed in the context of selection between pairs of stimulus locations. However, because selection between more than two stimuli requires comparisons between all the possible stimulus pairs, the results of this study apply directly to the general problem of selection among any number of competing stimuli.

3.7.1 Stimulus selection across space and model assumptions

Through computational modeling and subsequent experimental validation, our results showed that the multilobed RFs of sparse Imc neurons implement a combinatorial inhibition strategy for solving stimulus selection across locations (Figure 3.2). Our model included three key biological properties of the Imc-OT circuit as axiomatic features: (i) that Imc neurons are scarce (in elevation), (ii) that they interconnect with the OT in a specialized, spatially inverting manner, and (iii) that selection along elevation in the OT depends entirely on the priorities of the competing stimuli but not on their specific locations. The first axiomatic feature was demonstrated in Figure 2.6. The, second, although not yet confirmed functionally, is consistent with anatomical tracing studies [85]. Chapter 4 of my dissertation covers this question and confirms the presence of a functional spatially inverting connectivity in the Imc-OT circuit. The third, invariance of competitive suppression along elevation, has been demonstrated in previous work [75]. Incidentally, in other species, there are asymmetries between upper and lower hemifields [121] and, more generally, there can be a spatial gradient with stimuli at different locations being weighted differently; such a weighting function is easily incorporated into our optimization model.

As the goal of our modeling was to extract core computational principles underlying the implementation of selection at all location-pairs, we chose, in the interest of model simplicity, not to include other details of Imc RFs observed in the data Figure 2.5 beyond the three axiomatic features. These included: proportions of single versus multilobed RFs, the relative spreads of the heights of single vs. two- vs. three-lobed RFs, etc. In addition, the model made two other simplifying assumptions. First, to reduce model complexity, model RFs were implemented as discrete, binary 'pixels' (Figure 3.2), as opposed to being continuous hills of activity (Figure 2.2). As described in Figure 3.1, this assumption does not affect the spatial pattern of Imc inhibition onto the OT. However, it fails to capture the scaling of the strength of the inhibition based on the specific position of the stimulus within the half-max extent of a RF. Second, the model implemented the biological observation that the maximum number of lobes in any Imc RF (denoted k_{max}) was three (Figure 2.2L) by constraining the number of RF 'pixels' to be less than or equal to $k_{max.}$ The consequence of this assumption was that the sizes of model RFs were also limited to k_{max} pixels (Figure 3.3E, Figure 3.4B), unlike biological RFs that could be larger in size (Figure 2.2, Figure 2.1, Figure 3.5). This assumption was nonetheless necessary because it improved the convexity of the cost function, thereby allowing the runs of the optimization model to converge. Despite these assumptions, however, strong predictions of the model regarding combinatorially optimized inhibition for location-invariant selection (and the corresponding signature properties) were successfully validated with subsequent experiments, demonstrating that the Imc does indeed employ combinatorially optimized inhibition for stimulus selection at all possible elevational-pairs.

Notably, the signature properties were also robust to the specific proportions of single- vs. two- vs. three-lobed RFs. This was evidenced by the fact in the model, the signature properties held true for all optimal solutions even though the proportions of lobes varied widely across optimal solutions: for instance, one of the optimal solutions for L = 40, N* = 27 consisted of 7.5 % single-lobed, 18.5% two-lobed and 74% three-lobed RFs, whereas one optimal solution for L = 5, N* = 4 consisted of 50% single-lobed, 25% two-lobed and 25% three-lobed RFs (Figure 3.2A). Consequently, our conclusions were not impacted by a match (or not) between proportions of RFs in the model Imc and the proportions in the owl Imc.

(Incidentally, because the proportions in experimental data were obtained, necessarily, by combining data across coronal planes and across birds, they do not represent the distribution within any particular Imc coronal plane (Figure 2.2L).

3.7.2 Selection across locations in elevation vs. azimuth

Our results indicate that the Imc implements stimulus selection for locations both along elevation and azimuth. However, the strategy employed for location-pairs along the elevation is distinct from that along the azimuth, a difference that can be accounted for by the difference in anatomical organization along the two axes. The scarcity of Imc neurons in coronal planes is consistent with a combinatorially optimized solution for selection along elevational locations through the use of multilobed spatial RFs (Figure 2.6, Figure 3.2, Figure 3.3 and Figure 3.5). In contrast, abundance of Imc neurons along the rostrocaudal axis (section 2.3.2.2) is consistent with the 'modular copy-and-paste' solution through the use of single-lobed spatial RFs (Figure 2.1D).

3.7.3 Minimization of wiring and metabolic costs

In this context, three questions regarding the biological implementation of selection at all location-pairs by the Imc remain puzzling. First, why might $N < L$ be biologically desirable in the Imc in the first place, necessitating combinatorially optimized inhibition? Second, if $N < L$ is attractive biologically, why don't Imc RFs have a large number of lobes, thereby achieving greater savings in the number of Imc neurons (Figure 3.2A)? In other words, why is the maximum number of Imc RF lobes restricted to a low number ($k_{max} = 3$; Figure 2.2L)? Third, why is multilobed encoding found only along one spatial axis (here, elevation), why not along both axes for greater neuronal savings?

To gain insight into these questions, we examined Imc function in the context of two types of costs that nervous systems must incur in building and operating a neural circuit: wiring cost and metabolic cost. We estimated wiring cost by quantifying the cost of implementing spatially inverting projection patterns from the Imc to the OT [3], and metabolic cost by quantifying the cost of broadcasting of spikes across the OT for competitive suppression. We found that wiring cost decreases as the number of RF lobes increases (Figure 3.6A; section 3.7.3.1). In contrast, metabolic cost increases as the number of RF lobes increases (Figure 3.6B; section 3.7.3.1). Consequently, the wiring cost places a lower bound on the number of RF lobes (and a corresponding upper bound on the number of neurons), whereas the metabolic cost places an upper bound on the number of RF lobes (and a lower bound on the number of neurons). The optimal number of RF lobes (and the number of neurons necessary), therefore, is one that minimizes some weighted combination of the two opposing costs (Figure 3.6C). Because Imc neurons have high firing rates (median = 76.5 Hz [87, 108]; Figure 2.2A), this causes the metabolic cost of Imc function to scale up substantially, pulling the ideal number of RF lobes to even lower values than for low firing-rate neurons (Figure 3.6C, thick vs. thin line; thereby also providing a rationale for the continued presence of some single-lobe neurons in the Imc; Figure 2.2L).

Taken together, these results indicate that a small number of Imc neurons ($N < L$), with multilobed RFs that have a small number of RF lobes (small k_{max} value), are well suited to achieve selection across all locations if net neural costs are to be minimized. Increasing excessively the number of RF lobes along one spatial axis (here, elevation), or increasing the number of RF lobes also along the other axis as well (here, azimuth) are not biologically desirable. Thus, the estimation of the net cost of neural circuit operation provides a plausible window into 'why' the owl Imc may be organized functionally in the way that it is (Figure 3.6).

The reason for occurrence of multilobed encoding along elevation, specifically, rather than along azimuth, is less evident, and may simply be a consequence of the relative anatomy of Imc and OT. The Imc is asymmetric in shape, elongated along the rostrocaudal axis, but compressed along the dorsoventral axis [85]. It is plausible that the OT's representation of azimuth along its rostrocaudal axis [77] drives azimuths to be encoded along the parallel (and neurally rich) rostrocaudal axis of the Imc [85], for ease of wiring, thereby relegating elevation to be coded by the transverse (and neurally sparse) planes.

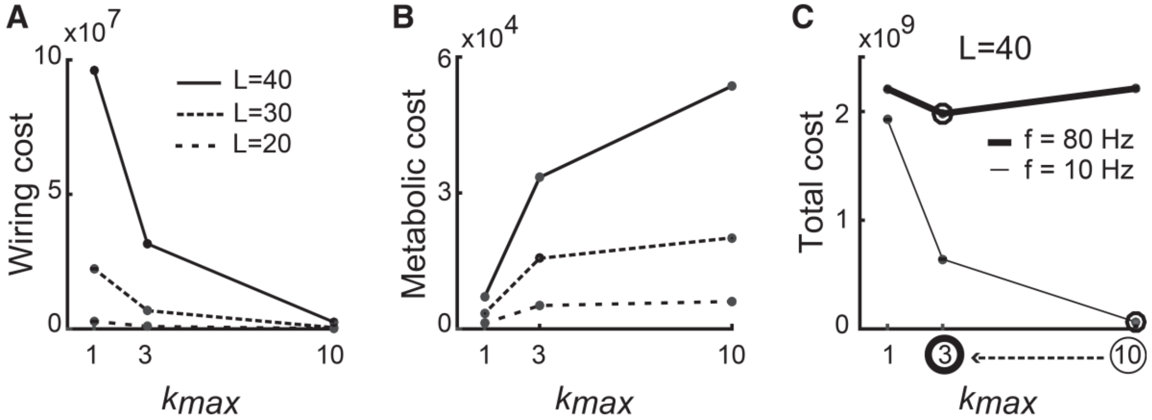


Figure 3.6: Metabolic and wiring costs for stimulus selection at all possible location-pairs.

(A) Wiring cost plotted as a function of the maximum number of Imc RF lobes allowed (kmax); calculated across optimal model solutions (section 3.7.3.1). Values of kmax examined along x-axis are particularly relevant to the Imc: kmax=1 corresponds the single-lobed case, kmax=3 to the experimentally determined value in the owl Imc, and kmax=10 to the practical upper bound on the number of possible RF lobes (based on the functional properties of Imc neurons; section 3.7.3.1 and Figure 2.2L). (B) Metabolic cost as a function of kmax (section 3.7.3.1). (C) Schematic showing total cost (weighted combination of (A) and (B)) for Imc circuit to solve selection at all location-pairs for L = 40 at low average firing rates (thin line: 10 Hz), and high average firing rates (thick line: 80 Hz; section 3.7.3.1); weights used for combining wiring and metabolic costs were identical for the low and high firing rate cases. Circled values along x-axis (and corresponding large dots) indicate the optimal kmax values at the two firing rate levels. Results demonstrate left shift of optimal kmax with increasing firing rates (section 3.7.3.1; the specific values of optimal kmax yielded by this analysis, and shown here, are incidental to the choice of the values of relative weights, and not particularly informative by themselves.) In all cases: mean ± SD values are plotted; SD values smaller than size of dots.

3.7.3.1 Calculating wiring and metabolic costs

Wiring cost: The wiring cost incurred by the Imc to implement selection at all location-pairs is estimated as the cost of generating axonal projections (‘wires’) between each Imc neuron and each of its target OTid neurons. This cost depends both on the number of locations that each neuron must suppress and the number of neurons in the population. Assuming that the lengths of wires from Imc to each OTid neuron is approximately equal (say 1 unit each, without loss of generality), we can estimate the total wiring length and consequently the total wiring cost using Eqn. 3.10 below (see [3]).

$$Wiring\ Cost(L, N^*, \qquad k_{max}) = (\sum_{i=1}^{N*}(\text{\#Locations suppressed by neuron } i\ ))\hat{}p \qquad - \quad (3.10)$$

The summation is the total wiring length of all the wires from the Imc neurons to the OTid population. ‘*p*’ is a power term such that typically $1 < p < 4$ (see [3]). This quantity is computed for each optimal solution (obtained over the 1000 runs) for a given (L, N*, k_{max}) triplet, and the results are plotted in Figure 3.6A.

Metabolic cost: The metabolic cost incurred by the Imc to implement selection at all location-pairs is estimated as the cost of generating and broadcasting spikes to the OTid to achieve competitive suppression. This depends on the number of neurons activated by a stimulus at each of the L locations, as well as the number of OTid locations to which each activated neuron delivers inhibition. If the cost of suppressing one OTid location using 1 spike is 1 unit, then the total metabolic cost for the circuit for a given firing rate *f* is given by Eqn. (17) below (using a similar formula as for wiring cost).

$$Metabolic\ cost(L, N^*, k_{max}, f)$$

$$= (\frac{f}{L}\sum_{j=1}^{L}\sum_{i=1}^{N*}(\text{\#Locations suppressed by neuron } i \text{ when stimulus is placed at location } j))\text{^q}$$

- *(3.11)*

Note that when the stimulus is placed at location *j*, the term in the inner summation is non-zero only for activated neurons. *'q'* is a power term chosen such that 1 < q < 4 (similar to the wiring cost). This quantity is computed for each optimal solution (obtained over the 1000 model runs) for a given (*L, N*, k_{max}, f = 10 Hz*), and the results plotted in Figure 3.6B.

Total cost: The total cost for any solution is calculated as a weighted combination of the wiring cost (weight = α) and the metabolic cost (weight = β) as given in Eqn. (18) below.

$$Total\ cost\ (L, N^*, k_{max}, f) =$$

$$\left(\alpha * Wiring\ cost(L, N^*, k_{max})\right) + \ (\beta * Metabolic\ cost(L, N^*, k_{max}, f))$$

- *(3.12)*

Therefore, there are five parameters (α, p, β, q and *f)* in the computation of the total cost. The results are plotted for $\alpha = 20,\ p = 2.5,\ \beta = 80,\ q = 2.42$ and for firing rates of *f* = 10 Hz (thin black line) and *f* = 80 Hz (thick black line) in Figure 3.6C.

4. Neural circuit mechanism for generating categorical representations

4.1 Introduction

Selection among competing options is a fundamental component of adaptive behavior. Categorical neural representations have been reported in multiple brain areas to underlie selection and decision-making across animal species [94, 95, 122-129]. These representations, which involve the transformation of continuously varying inputs into discrete output classes, involve a large, abrupt change in responsiveness across the selection boundary (Figure 4.1A – red). They enhance the reliability of neural signaling of the selected category (or 'winner') specifically when competing options on opposite sides of the selection boundary are very similar to one another (i.e., when there is input ambiguity), and in the face of neural response variability (i.e., neural 'noise') (Figure 4.1A-red vs. orange).

This overall robustness to sensory and neural noise, of the decoding of the winner from neural responses, can be quantified using boundary discriminability (*bd'*) - the discriminability between two competing options on either side of the selection boundary (d-prime; section 4.7.2): bd' is higher for more categorical response profiles (Figure 4.1A and Figure 4.1B-left, open symbols: red vs. orange). However, *bd'* is also higher for response profiles that are more robust-to-noise *without* being more categorical - for instance, scaled response profiles (Figure 4.1A- orange vs blue). By contrast, a metric that is insensitive to scaling and that captures, specifically, the strength of categorization by response profiles is the commonly used categorization index ([95, 123]; it operates effectively on the normalized response profiles; Figure 4.1B-middle inset). A modified version, *CatI,* takes into account response variability as well: it compares the average between-category discriminability with the average within-category discriminability (Figure 4.1B-right, filled symbols; section 4.7.3). Thus, both scaling-up of responses as well as enhancing their categorical nature can improve the overall robustness-to-noise of the neural signaling of the winner (Figure 4.1AB, black vs. red). Biologically, however, the limited dynamic ranges of neural responses (which, in turn, limit the range of scaling possible), as well as the inherent discreteness of categorical responses, make categorical neural representations a preferable solution to promote robust-to-noise selection.

Despite the clear utility of categorical representations, their pervasiveness in brain areas [94, 95, 122-129], as well as the range of theoretical formulations and computational models proposed to account for them [130-133], identifiable neural circuit mechanisms that control their production have remained elusive. Specifically, it is unclear what aspects of circuit architecture are essential for generating categorical responses.

4.2 Donut-like inhibitory motif controls categorical signaling in model circuits

We reasoned that properties of neural inhibition, which is known to be involved in comparing representations of competing options [75, 134-136], would impact the nature of competitive response profiles. This led to two concrete hypotheses re circuit mechanisms. The first was that feedback inhibition between the options, i.e., inhibition whose strength is iteratively influenced by both options, may play a key role. This was motivated by extensive modeling of categorical decision-making [132, 137], work on direction selectivity in the retina [138], as well as work on spatial selection in barn owls [139]. The second hypothesis was that a donut-like pattern of competitive inhibition, i.e., one in which each option suppresses others more strongly than it suppresses itself, may play a key role. This was motivated by work in turtles [99] as well as by the need, in categorical profiles, for enhanced response differences across the selection boundary: we reasoned that having strong inhibition to 'other' options, but weak "self" inhibition may enhance response differences.

To assess the computational validity as well as relative efficacies of these hypothesized circuit mechanisms, we turned to modeling. Starting with a generic baseline circuit capable of comparing the representations of competing options (Figure 4.1C- model on top-left; section 4.7.2;[134-136]), we examined the impact of introducing just feedback inhibition (Figure 4.1C- bottom-left), just donut-like inhibition (Figure 4.1C- top-right), or both (Figure 4.1C- bottom-right), on the production of categorical response profiles in the presence of noise. Response profiles were measured to a classic two-stimulus morphing protocol in which the relative strength of the two stimuli (S1 and S2) was systematically varied, resulting in two categories (S1> S2 and S1<S2; Figure 4.1E-bottom, graphic; [95, 123, 125]). Model neurons in the circuit

were simulated with noisy, sigmoidal input-output functions (fano factor of 6; 30 repetitions per 'neuron', n=50 'neurons'; section 4.7.2). The resulting response profiles from the four models clearly involved a scaling component (Figure 4.1E-left**)**, and therefore, we compared both scaling-dependent robustness-to-noise (*bd'*) as well as the scaling-insensitive categorization strength (*CatI)* across the response profiles (section 4.7.3).

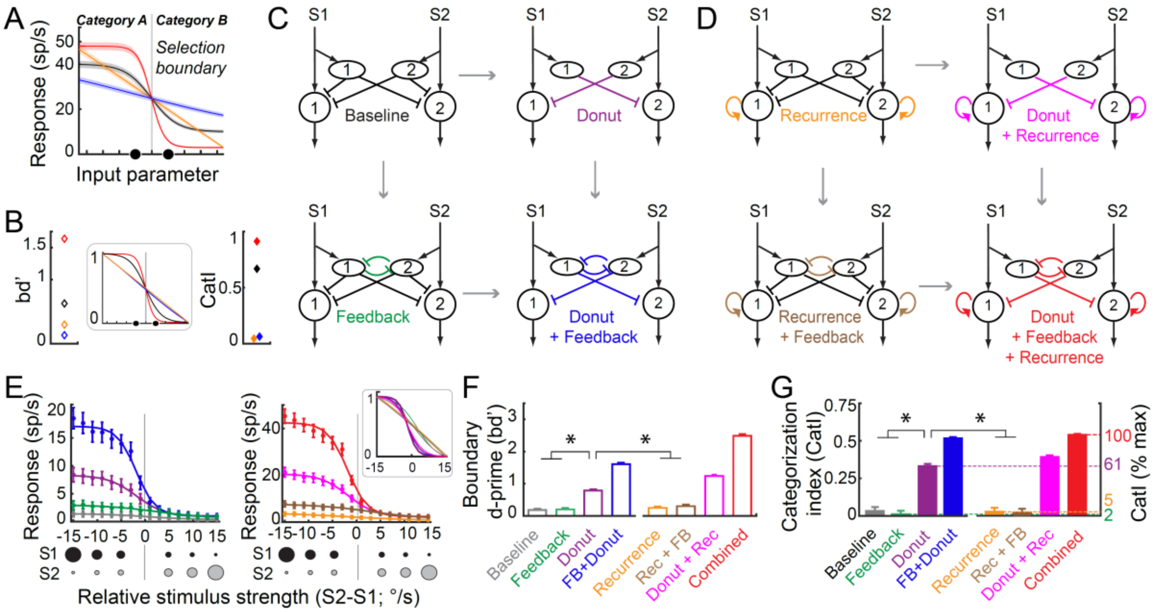


Figure 4.1: Donut-like inhibition surpasses other hypothesized circuit motifs in generating categorical responses to competing stimuli.

(A) Schematic showing 4 different mathematically generated response profiles, as a function of continuously varying input. Vertical line: categorization or selection boundary. Translucent band: variability in responses; fano factor of 6 used to generate these responses (section 4.7.2). Black filled dots: Two competing inputs just straddling the selection boundary: nearly equal in input value (at a distance of 3 units from the selection boundary), but belonging to different categories. (B) Left (open symbols): Boundary discriminability (bd'); defined as d-prime between the responses to the two competing inputs indicated in A (section 4.7.3), computed for each response profile in A (colors). Middle-inset: Response profiles in A normalized between 0 and 1; only means are shown; red profile transitions more abruptly than black, orange and blue profiles are equally gradual (linear). Right (filled symbols): Categorization index (CatI; accounts for response variability as well; section 4.7.3) for response profiles in A. [95, 123]. (C, D) Computational models incorporating different circuit motifs (section 4.7.2; Figure 4.2). All models built upon generic baseline circuit capable of comparing competing options (C; top-left). Each model shows two 'channels'; each channel is the group of neurons (numbered) involved in representing a stimulus (S1 or S2); the goal of these models is to signal if S1>S2 (category 'A') or S2>S1 (category 'B'). Circles - output neurons; ovals - inhibitory neurons; arrows with pointed heads - excitatory connections, arrows with flat heads - inhibitory connections. (C) Four circuit models: baseline circuit (top-left), inclusion of feedback inhibition (bottom-left), donut-like inhibition (top-right) or both (bottom-right; section 4.7.2). (D) Four more circuit models: baseline circuit with recurrent amplification in each channel (top-left), followed by inclusion of feedback inhibition (bottom-left), or donut-like inhibition (top-right) or both (bottom-right). (E) Simulated response profiles of 'example' output neuron 1 from each of the models in C (left) and D (right); obtained using a two-stimulus (S1 and S2) morphing protocol (bottom inset; section 4.7.4). Responses are mean ± s.e.m of 30 repetitions. The continuously varied input parameter was the relative strength of the two stimuli (S2-S1). Lines – best sigmoidal fits. Bottom schematic: Morphing protocol; stimulus strength schematized by size of dot. Input-output functions of model neurons were sigmoids with Gaussian noise (section 4.7.2; fano factor=6). Right-Inset: Response profiles normalized between 0 and 1 (same conventions as B-inset). (F, G) 'Population' summary of bd' (F) and CatI (G) of response profiles from various circuit models (colors); mean ± s.e.m.; n=50 model neurons. '*': $p<0.05$, ANOVA followed by Holm-Bonferroni correction for multiple paired comparisons; only a key subset of significant differences indicated for clarity. See also Figure 4.2.

We found that feedback inhibition between competing channels had no effect on *bd'* (Figure 4.1EF – green vs. grey; p= 0.99) or on *CatI* (Figure 4.1E-inset and G – green vs. grey; p =0.99), and this result was largely independent of the strength of feedback inhibition (Figure 4.2A). By contrast, introduction of a donut-like pattern of inhibition provided a boost in robustness-to-noise (Figure 4.1EF – purple vs. grey; p =5.99 e-8), and an even greater boost in the strength of categorization of the response profiles (Figure 4.1E-inset and G – purple vs. grey; p = 5.98e-8). The magnitudes of improvement were inversely related to the strength of "self" inhibition, reaching the maximum when "self" inhibition was zero (Figure 4.3B; *bd'*: $\varrho = -0.89, p = 1.64$ *e-4*; *CatI*: $\varrho=-0.81, p = 2.6$ *e-3,* Pearson correlation test), thereby revealing the importance of donut-like inhibition over feedback inhibition for categorical responses. These results held true also when we started with a circuit that included recurrent excitation, a common element in models of decision-making thought to play a role in categorical selection (Figure 4.1D, and E-G: right panels, Figure 4.2C; [132, 140, 141]).

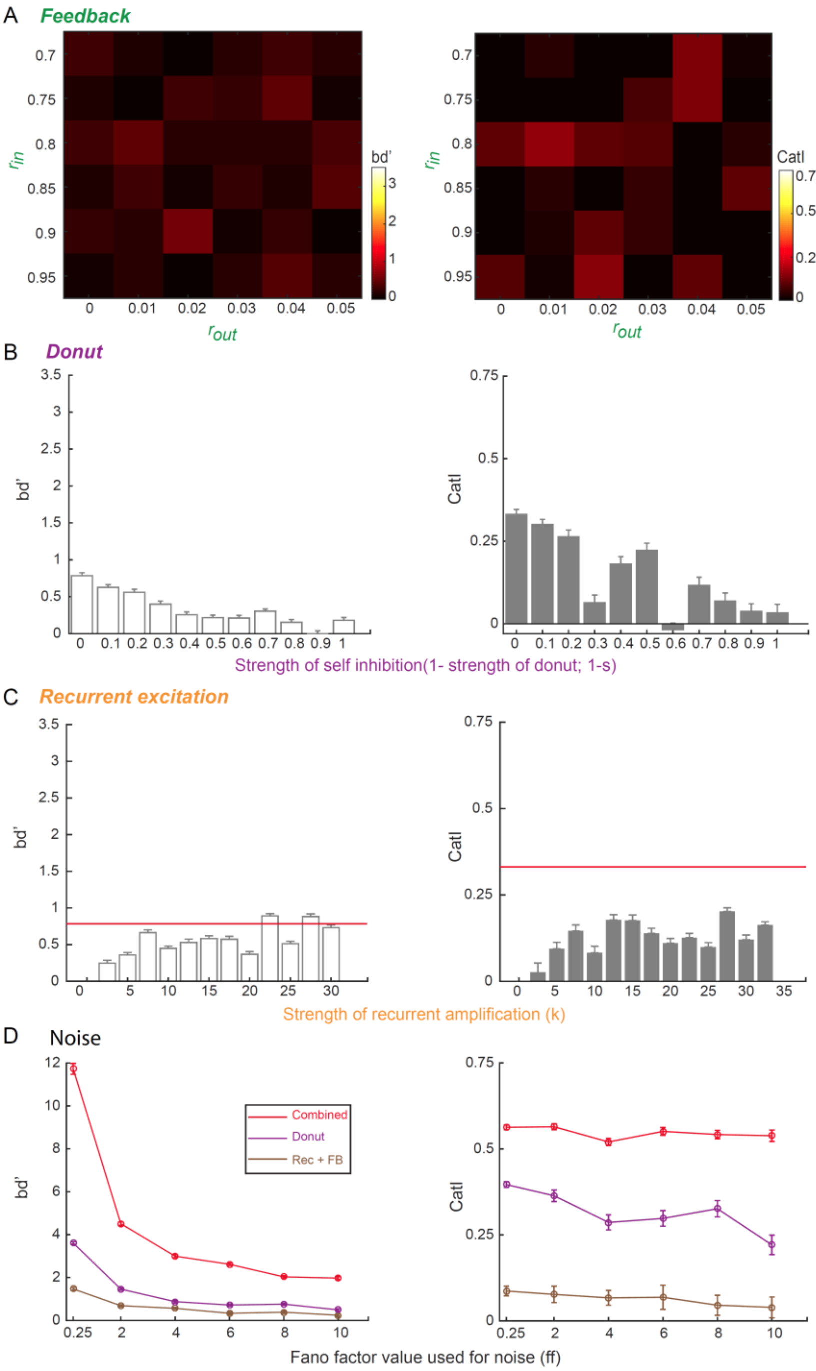


Figure 4.2: Donut-like inhibition surpasses other circuit motifs over a large range of values of key model parameters and amounts of response noise.

(A) Effect of varying strength of feedback inhibition in model in Figure 4.1C-bottom-left, on bd' (left) and CatI (right). r_{in}: input divisive factor; r_{out}: output divisive factor (section 4.7.2). Range of variation based on previously published work [139]. Maximum bd' over the range of parameters is = 0.5478 (lower than from the donut-like motif, bd' = 0.785); maximum CatI = 0.157 (lower than from the donut-like motif, CatI = 0.331). **(B)** Effect of varying strength of "self" inhibition in model in Figure 4.1C-top-right on bd' (left) and CatI (right). Left: Corr = -0.89, p = 1.64e-4 (correlation test). Right: Corr = -0.81, p = 3e-

3 (correlation test). Maximal effects on bd' and CatI are when "self" inhibition=0, i.e., when the circuit has donut-like inhibition. **(C)** Effect of varying strength of recurrent amplification in model in Figure 4.1D-top-left, on bd' (left) and CatI (right). Red lines: values from circuit with donut-like motif only. Left: bd' increases with amplification (corr=0.68; p = 1e-3; corr test), saturating around k=10; the increase is as expected because of the progressive scaling up of responses due to amplification. Note that physiological range is under 3 [142]. Right: However, CatI does not change systematically (corr= -0.14; p =0.54; corr test), indicating that varying the strength of amplification either does not affect the estimates in Figure 4.1F for this motif. **(D)** Comparison across three models of bd' (left) and CatI (right) from simulated response profiles as a function of fano-factor. Models are circuit with donut-like inhibitory motif only (Figure 4.1C, top-right), or circuit with recurrent amplification and feedback inhibition together (Figure 4.1D, bottom-left), or circuit with all three motifs combined (Figure 4.1D, bottom-right).

Donut-like inhibition, therefore, emerged as the most important single motif for producing categorical responses (Figure 4.1G, right axis; 61% of combined *CatI*), and this held true over the range of fano factor values tested (Figure 4.2D). Whereas the presence of feedback inhibition or both recurrent excitation and feedback inhibition enhanced the impact of the donut-like motif (Figure 4.1FG: blue vs. purple, red vs. purple; p< 6.023 e-8 in both cases), without the donut-like motif, they were nearly ineffective, either individually or together, at signaling the winning option categorically in the presence of input ambiguity and neural noise (Figure 4.1FG: green, orange, brown).

4.3 Functionally donut-like inhibitory motif operates in the barn owl midbrain

We wondered, next, if the donut-like inhibitory motif operates within a neural selection circuit. An excellent site in the brain to investigate this is the midbrain spatial selection network in vertebrates [79, 143]. It includes the superior colliculus, SC (optic tectum, OT, in non-mammals), a sensorimotor hub that is required for target selection for spatial attention [92, 93]. Neurons in the intermediate and deep layers of the SC (SCid/OTid) encode a topographic spatial map of stimulus priority (priority = physical salience + behavioral relevance) [59], and OTid neurons signal the highest priority stimulus categorically*:* they respond with a high firing rate when the stimulus inside their RF is the one with highest priority, but their responses are suppressed to a low level when the RF stimulus is no longer the one with highest priority [94, 95, 144]. The source of this competitive response suppression in the OTid is a group of GABAergic midbrain tegmental neurons called Imc in birds (isthmi pars magnocellularis; [85]): Imc inactivation abolishes all competitive interactions in the OTid and thereby its ability to signal the winner [87, 100].

This network has been shown to contain both feedback inhibition among competing channels and recurrent excitation within each channel. It implements the former as direct, long-range inhibition between Imc neurons (Figure 4.3A-right – blue; not shown in 2A-left for clarity; [85, 145]), and the latter, as response amplification by a group of cholinergic neurons, Ipc (Figure 4.3A; orange), that connect in a point-to-point manner with the OT [86, 88, 142].

What, however, of the donut-like motif? Imc neurons are known to receive input from a focal portion of the OT (within layer 10; OT10), but to suppress OTid activity broadly across the space map by sending projections via two pathways (Figure 4.3A): (a) directly to the OTid (layers 11-15) (Figure 4.3A – purple;[85]), and (b) indirectly to the OTid by inhibiting the potent point-to-point amplifiers of OTid activity, namely the cholinergic Ipc (Figure 4.3A –orange; [85, 87]). Incidentally, simulations with a circuit model in which the recurrent amplifier is also the recipient of powerful inhibition, as in the OT-Imc-Ipc circuit, confirm the superiority of the donut-like motif for producing categorical representations (Figure 4.4A-C).

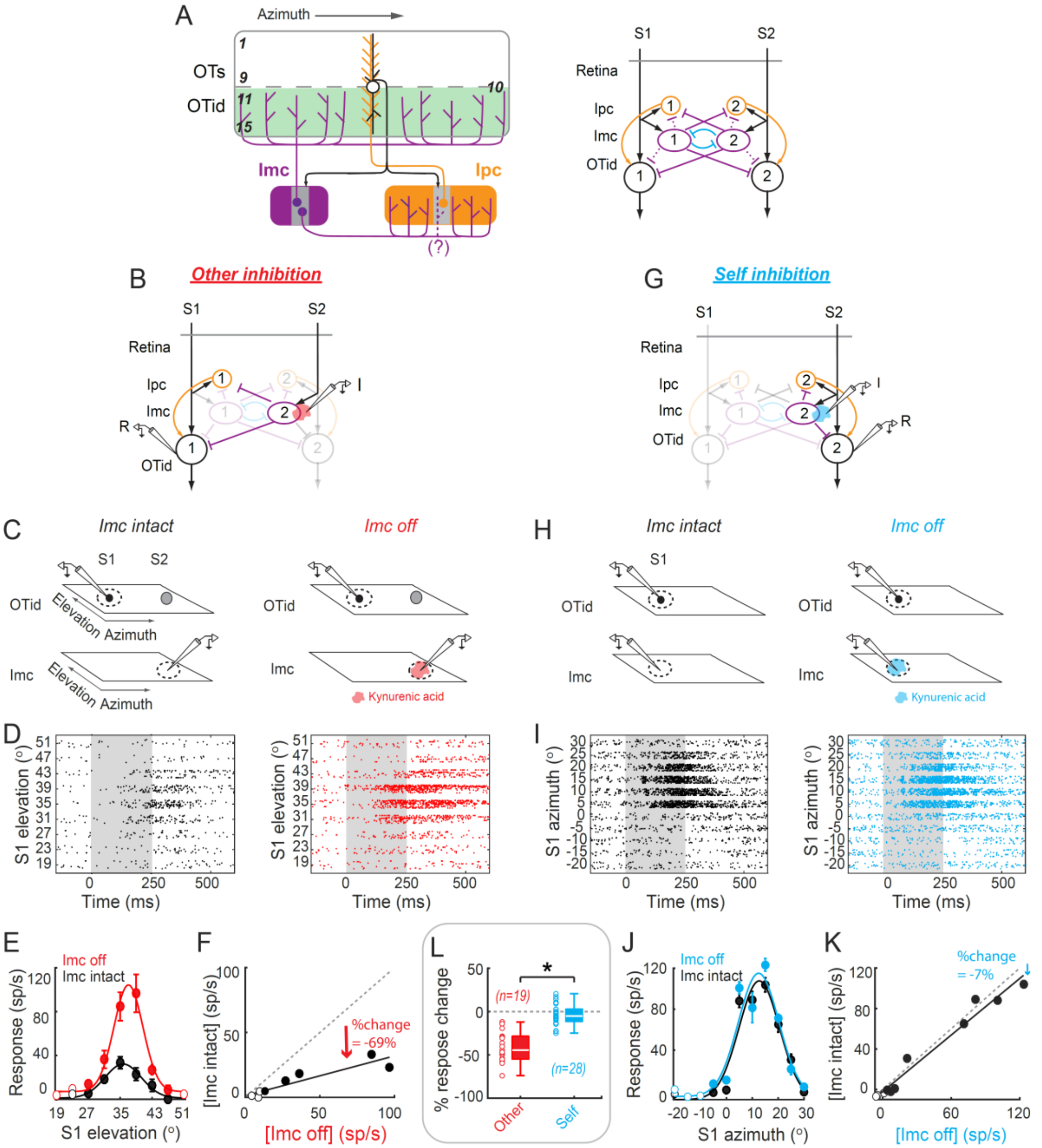


Figure 4.3: Barn owl midbrain selection network contains functional donut-like inhibitory motif operating along azimuthal space.

(A) Schematic of avian midbrain selection network. Left: Superficial (OTs), and intermediate-deep (OTid) layers of optic tectum; numbers: individual layers (1-15). Columns across layers of OT tissue, from left to right, encode individual locations in space (here, azimuth) topographically [77]. Imc (purple): Isthmi pars magnocellularis, inhibitory neurons in midbrain tegmentum. Ipc (orange): Isthmi pars parvocellularis, cholinergic neurons. Imc and Ipc receive focal input from OT10 (dark grey projections). Ipc projects back focally to OT (orange projections). Imc projects back broadly across OTid and Ipc (purple projections); thought to spare the portion of OT providing input (absence of purple projections in the column occupied by black OT10 neuron[85]; '?': unknown if congruent portion of Ipc space map is also spared of Imc projections. Right: Network model showing the OT-Imc-Ipc circuit; conventions as in Figure 4.1. Simulations from this model in Figure 4.4. Dashed lines: Unknown if these connections exist functionally. (B-F) Measurement of the strength of net "other" inhibition from Imc → OTid with paired recordings in Imc and OTid. "Net" indicates the combined inhibition due to both the direct (Imc→ OTid) and indirect (Imc→ Ipc→ OTid) pathways (text). "Other" indicates that the OTid neuron encodes for (distant) spatial locations outside Imc neuron's RF. (B) Experimental setup. Iontophoresis and recording electrode (I) in the portion of Imc (encoding for stimulus S2); recording electrode (R) in the portion of OTid encoding for distant location (and stimulus S1). (C) Schematic of OTid and Imc space maps (quadrilaterals) showing RFs of neurons being recorded (dotted ovals) and stimulus protocol (black filled dot – S1; gray filled dot – S2). S1 and S2 are looming visual stimuli of fixed contrast but different loom speeds (strength; Mysore 2010); S1 = 9.6 °/s, S2 = 19.2°/s (section 4.7.4). (D) Raster responses of example OTid neuron to paired stimulus protocol in C, in the Imc-intact condition (left column; black data) and Imc-off condition (right column – red data). Imc inactivation by (reversible) iontophoresis of kynurenic acid (a pan-glutamate receptor blocker;[100]; 4.7.4). Gray shading: stimulus duration (E) Response firing rates of this OTid (computed from rasters in D; spike count window = 100-350 ms); mean ± s.e.m. Lines: best Gaussian fits. Filled dots: responses to S1 at locations inside the OTid RF; open circles, outside RF (section 4.7.4). (F) Scatter plot showing OTid responses in Imc-intact vs. Imc-off conditions. Line: Best fit straight line; slope = 0.31, r^2 =0.76. % change in responses = 100* (responses in Imc intact condition – responses in Imc off condition)/responses in Imc off condition = 100*(slope-1), directly estimates the strength of inhibition at this OTid neuron due to Imc (here, -69%; 100*(0.31-1); section 4.7.4). (G-K) Measurement of the strength of net "self" inhibition from Imc → OTid with paired recordings in Imc and OTid. (conventions as in B-F). (G, H) OTid neuron is spatially aligned with Imc neuron (both encode overlapping locations); distance between OTid and Imc RF centers = 1.5°. (K) Line: Best fit straight line; slope = 0.93, R^2 =0.95. % change in responses directly estimates the strength of suppression at this OTid neuron due to Imc (here, -7%; section 4.7.4). (L) Population summary of strength of net "other" inhibition (red; n=19 Imc-OTid pairs), and strength of net "self" inhibition (blue; n=28 pairs) from Imc→ OTid. Average distance between OTid and Imc RF centers in "other" experiments = 26.8° +/- 2.3°; in "self" experiments = 2.9° +/- 0.7 °. $p<1e-5$ (red vs. blue), $p=<1e-5$ (red vs 0), $p=0.12$ (blue vs. 0), paired t-tests with HBMC correction (section 4.7.4). See also Figure 4.4.

Whereas anatomical tracing studies [85] have indicated that the direct projections from the Imc to the OT may spare the portion of the OT providing input to the Imc (Figure 4.3A; absence of purple projections in portion of OT around black neuron), this not been established functionally. Crucially, whether or not the indirect inhibitory pathway involving the Ipc exhibits the donut-like motif is unknown (Figure 4.3A, purple '?'). This is critical because this indirect pathway is known to constitute the dominant route of inhibition from

Imc to OTid [85]: the majority of Imc projections target the Ipc (rather than the OTid; [85]), and Ipc provides substantial amplification of OTid activity (1.43x, on average; [142]).

To address functionally whether the Imc-Ipc-OT circuit implements a donut-like pattern of competitive inhibition, we made dual extracellular recordings in the barn owl OTid and Imc (section 4.7.4; Figure 4.3BG). We measured directly the strength of the net competitive inhibition delivered by Imc neurons onto OTid neurons encoding for stimuli at distant locations ("other" inhibition), and separately, the strength of net inhibition delivered by Imc onto OTid neurons encoding for overlapping locations ("self" inhibition), (Figure 4.3BG).

4.3.1 "Other" inhibition

To measure the strength of "other" inhibition, we first recorded the responses of OTid neurons to a stimulus (S1) inside their receptive field (RF; Figure 4.3B, C-left; section 4.7.4) while simultaneously presenting a competing stimulus outside the RF (S2; at a distant azimuthal location from S1; section 4.7.4). Responses to S1 in the OTid are known to be divisively suppressed by a distant S2, by an amount depending on their relative strength [75, 107]. Notably, this suppression is known to be abolished upon focally inactivating the portion of Imc representing S2, revealing Imc as the primary source of competitive suppression in the OTid [87, 100]. Therefore, we estimated the strength of net "other" inhibition by comparing the responses to the paired presentation of S1 and S2 when Imc was intact (Figure 4.3C-left) versus when the portion of Imc encoding S2 (spatially mismatched with the OTid recording site), was focally and reversibly inactivated (Figure 4.3C-right). Any increase in OTid responses following Imc inactivation would represent inhibition provided by Imc. We quantified the strength of net "other" inhibition as: % change in OTid responses = 100* (responses in Imc intact condition – responses in Imc off condition)/responses in Imc off condition (section 4.7.4; no change in OTid responses would indicate zero competitive suppression by Imc onto that OTid neuron.)

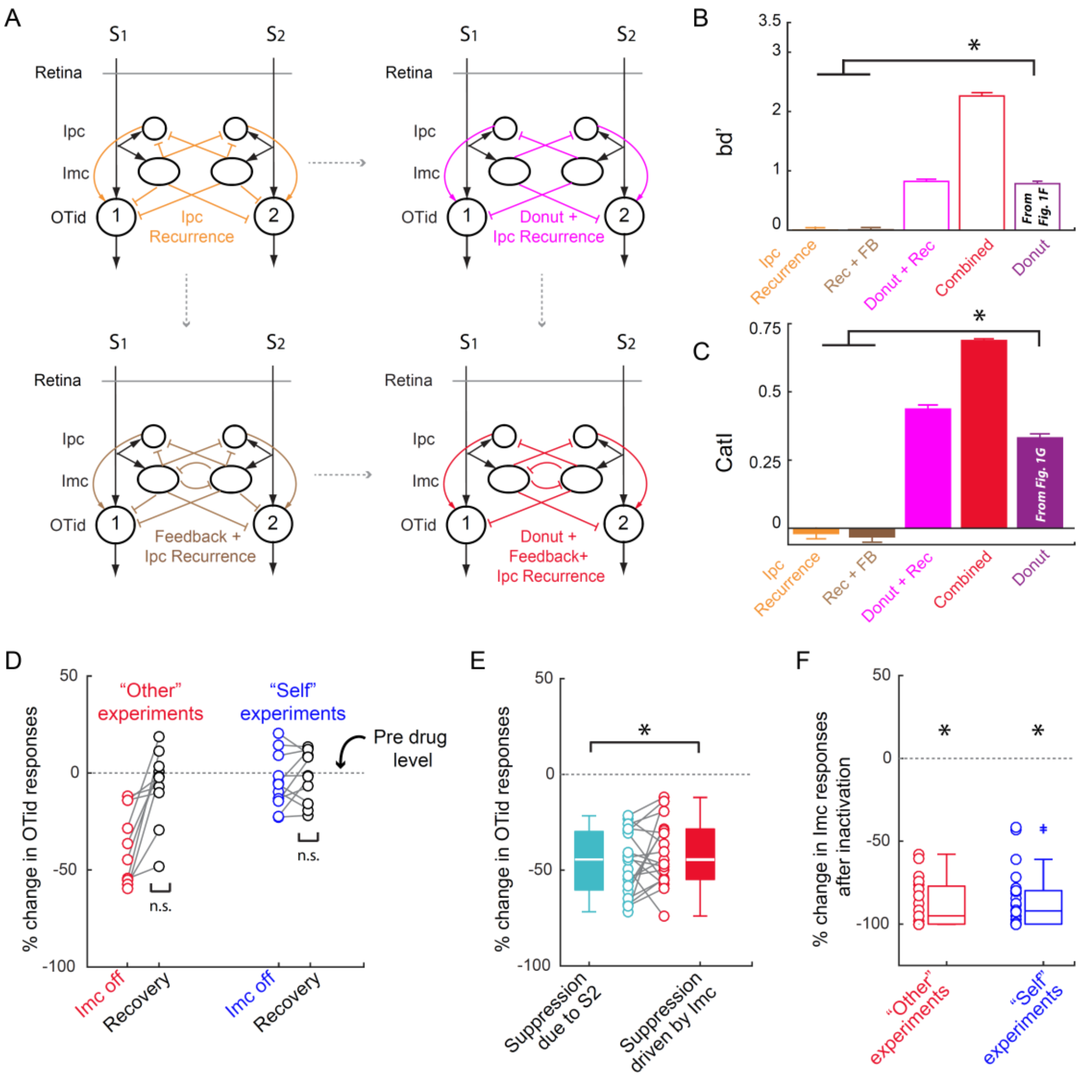


Figure 4.4: Donut-like inhibition surpasses other circuit motifs even in the case of biologically grounded model (model based on barn owl Imc-Ipc-OT connectivity and function), and supporting data for Figure 4.3 (experiments).
(A) Baseline + recurrent excitation model, similar to the one in Figure 4.1D, top-left, but with Ipc providing recurrent amplification, and receiving inhibition from Imc. Bottom-left: Adding feedback inhibition to model in top-left. Top-right: Adding donut-like motif to model in top-left. The donut-like motif is 'added' by setting "self" inhibition to zero; this is done both for the direct inhibitory pathway from Imc→ OTid, and the indirect pathway: Imc→ Ipc→ OTid. Bottom-right: Adding both feedback inhibition and donut-like inhibition. **(B)** bd' for these four models in A, computed on the response profiles of model neuron 1 obtained by simulating with the same two-stimulus morphing protocol used in

Figure 4.1E. Shown ghosted in for comparison are the results from the circuit model in Figure 4.1C, top-right, that has just the donut-like motif without any recurrent amplification or feedback inhibition. '*': p = 9.5e-10, ANOVA followed by HBMC correction for multiple comparisons. **(C)** Same as B, but CatI. Results show that the supremacy of the donut-like motif over recurrent amplification alone, or the combination of that with feedback inhibition, in generating categorical response profiles still holds in the (biologically grounded) model of the barn owl midbrain circuit. '*': p = 9.6e-10, ANOVA followed by HBMC correction for multiple comparisons. **(D)** Recovery of OTid responses from kynurenic acid iontophoresis for experiments in Figure 4.3. Data on left: "Other" experiments. OTid responses revealed significant change during iontophoresis (red dots; subset of the data reproduced from Figure 4.3L), but returned to pre-drug baseline (horizontal line) in recovery (black dots; p=0.32, t-test against 0). Recovery data obtained 15 min after iontophoretic eject current was switched to retain current; section 4.7.4). Data show recovery, demonstrating that the effects reported Figure 4.3 are due specifically to drug iontophoresis/Imc inactivation. Data on right: "Self" experiments. OTid responses showed no significant change during iontophoresis (blue dots; reproduced from Figure 4.3L), and stayed around zero in recovery (black dots; p=0.58, t-test against 0. **(E)** "Other" experiment. Comparison of suppression provided by Imc with that due to S2 (i.e., the maximum amount of suppression experienced by the OTid neuron in this stimulus protocol). This is done by comparing (i) % suppression of OTid responses by stimulus S2 when Imc is intact, computed by comparing OTid responses to S1 alone versus the paired presentation of S1 and S2 (teal;[75]), with (ii) % suppression of OTid responses to paired presentation of S1 and S2 produced by Imc inactivation, computed by comparing OTid responses in the Imc-intact to Imc-off conditions (red; data reproduced from Figure 4.3L). Consistent with [100], nearly all the suppression due to competitor S2 (teal) is supplied by Imc (red) (p= 0.22, t-test, teal vs. red), verifying that Imc supplies powerful "other" inhibition. **(F)** Quantifying the effectiveness of Imc inactivation by kynurenic acid iontophoresis in the "other" (red) and "self"-inhibition (blue) experiments. In both cases, the inactivation was highly effective. Red: median % change in response; median = 95%, 95% CI = [87%, 103%], p = 3.8 e-6, sign test; Blue: median % change in responses median =0.92, 95% CI = [86%, 98%], p = 7.45e-9, sign test.

We found that Imc neurons exerted strong inhibition onto OTid neurons encoding for distant, non-matched spatial locations (Figure 4.3F, L-red; mean strength = - 40.47 % +/- 17.70 %, n==19 paired neurons; p=9.43e-9, t-test against 0; mean distance between centers = 26.74 °). We verified that the results were specifically due to Imc inactivation by observing that OTid responses to paired S1 and S2 after recovery from iontophoresis (measured 15 min after the drug was turned off; section 4.7.4) returned to pre-drug levels (Figure 4.4D). Indeed, the suppression provided by Imc accounted for nearly all the suppression exerted by the stimulus S2 (Figure 4.4E). In these experiments, Imc was inactivated effectively (median = 95%, 95% CI = [87%, 103%]; p = 3.8e-6, sign test, n=19; Figure 4.4F).

4.3.2 "Self" inhibition

To measure the strength of self-inhibition, we recorded the responses of OTid neurons to a single stimulus (S1) presented inside their RF (Figure 4.3G; section 4.7.4). We compared these OTid responses when Imc was intact (Figure 4.3H-left) versus when the portion of Imc encoding S1, i.e., spatially matched to the OTid recording site, was focally and reversibly inactivated (Figure 4.3H-right). Any increase in OTid responses between the Imc-intact (Figure 4.3I-left; Figure 4.3J-black) and the Imc-inactivated (Figure 4.3-right; Figure 4.3-blue) conditions quantified the amount of suppression due to Imc onto the "self" OTid location, thereby directly estimating the strength of net 'self-inhibition': % change in OTid responses = 100* (responses in Imc intact condition – responses in Imc off condition)/responses in Imc off condition (section 4.7.4).

We found that Imc neurons exerted no significant inhibition onto OTid neurons encoding for overlapping spatial locations (Figure 4.3K, L-blue; mean strength = -3.7 %, s.d. = 12.2%, n = 28 neuron pairs; p = 0.12, t-test against 0; mean distance between centers = 2.86 °). We verified that these results were not due to ineffectiveness of iontophoresis by observing that the suppression of Imc responses by kynurenic acid was substantial (Figure 4.4F), and not distinguishable from that in the "other" case (Figure 4.4F; p = 0.68, ranksum test, Imc suppression by drug in "self" vs. "other" cases). Therefore, strength of net "self" inhibition in the OTid was substantially weaker than the strength of net competitive inhibition (Figure 4.3L, red vs. blue; p = 7.8e-11, two sample t test with HBMC correction; 4.7.4). Together, these findings demonstrated the presence of a (net) donut-like, spatial pattern of inhibition across the OTid azimuthal space map (i.e. when S1 and S2 were separated along the azimuth).

4.4 Donut-like inhibitory motif in the barn owl midbrain is multi-holed

Imc neurons are strikingly asymmetric in their encoding of elevational versus azimuthal space. The majority (67%) exhibit RFs with multiple discrete firing fields ("lobes") distributed along the elevation, but not the azimuth (multilobed RFs exhibit upto 3 lobes along the elevation; [146]; Figure 4.5A,G-left panels).This unusual RF structure has been shown to be essential for Imc to achieve selection at all possible pairs of spatial locations in the face of scarcity of its neurons, and it does so does using a novel, combinatorially optimized inhibitory strategy [146].

A direct consequence of multilobed encoding of elevational space is that there are gaps between the lobes of an Imc RF, constituting locations that are outside that neuron's RF (Figure 4.5AB-left panels; light red bar). If the donut-like inhibitory motif is to operate also for (categorical) selection among competing stimuli separated along the elevation, then the spatial pattern of inhibition must respect a very strict condition: namely, that, multilobe Imc neurons ought to send strong competitive inhibition to OTid neurons encoding locations in the gaps between RF lobes, but weak or no inhibition to OTid neurons encoding locations within the other RF lobes of the same Imc neurons (weak self-inhibition), predicting a 'multi-holed' donut (Figure 4.6A).

4.4.1 "Gap" and "different lobe" inhibition

To test experimentally if this strict requirement holds true, we again made dual extracellular recordings in the OTid and Imc. We first recorded from an Imc neuron, mapped out its spatial RF, and applied previously published analyses to determine if it was a multilobed

RF (two-lobed RF in Figure 4.5AB – left panels and Figure 4.6B-D; three-lobed RF in Figure 4.5GH – left panels and Figure 4.6E-G; section 4.7.4;[146]). Next, we positioned a second electrode in the OTid such that the spatial RF of the OTid neuron was either centered within the gap between Imc RF lobes (Figure 4.5AB – right panels and Figure 4.5C-left, light red bar) or it overlapped one of the lobes of the Imc neuron's RF (Figure 4.5GH – right panels and Figure 4.5I – left, light blue bar). We, then, recorded the responses of the OTid neuron to a stimulus inside its RF (S1; Figure 4.5C-left and I-left; section 4.7.4) while simultaneously presenting a competing stimulus (S2) at a distant location along the elevational axis such that S2 was within a (different) lobe of the RF of the Imc (Figure 4.5C-left – S1 in the Imc RF gap – light red bar and S2 within an Imc RF lobe; Figure 4.5I-left – S1 within one Imc RF lobe and S2 within a different one – light blue bar). We compared OTid responses when Imc was intact (Figure 4.5CI-left panels) versus when the portion of Imc encoding S2 was focally and reversibly inactivated (Figure 4.5CI-right panels). As before, any observed response increases directly estimated, respectively, the strength of net inhibition exerted by the Imc neuron onto the OTid neuron encoding the gap location ("gap" inhibition), or the strength of net inhibition exerted by the Imc neuron due to a stimulus within one of its lobes, onto the OTid neuron encoding for a location within a different lobe of the Imc's RF ("different lobe" inhibition; 4.7.4).

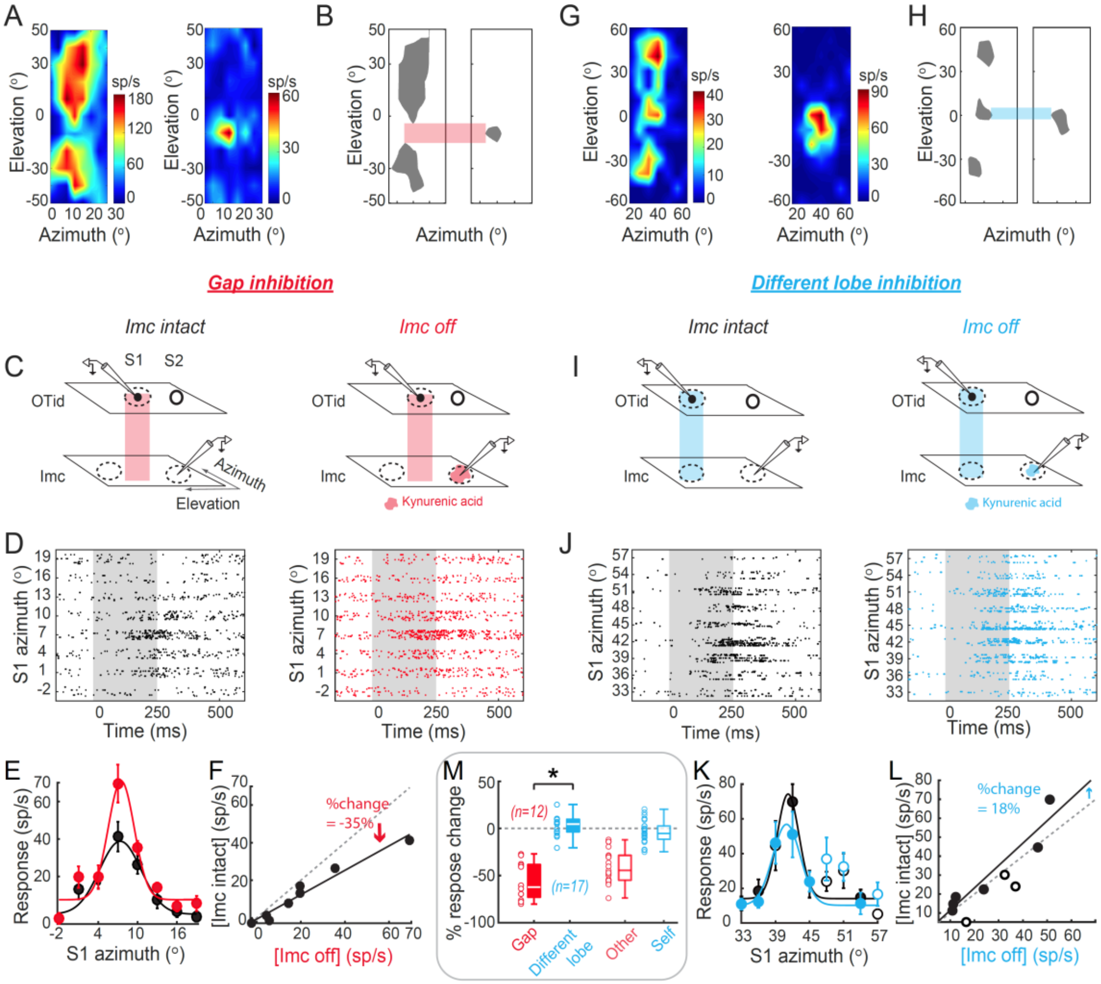


Figure 4.5: Midbrain selection network in barn owl contains multi-holed donut-like inhibitory motif operating across 2-D sensory space (azimuth x elevation).

(A-F) Measurement of the strength of net “gap” inhibition from Imc → OTid with paired recordings in Imc and OTid. (A) Spatial receptive fields (RFs) of an Imc neuron (left) and an OTid neuron (right) from a paired Imc-OTid recording experiment. Imc RF is two-lobed (text; Figure 4.6B-D; section 4.7.4; [146]). OTid RF lies in the gap between the lobes of the Imc RF. (B) Binarized versions of RFs in A, at 60% max. firing rate in each case. Red horizontal bar: highlights the relative position of OTid RF to Imc RF lobes. (C-F) Conventions as in Figure 4.3C-F. C, Red vertical bar indicates that OTid RF is in the gap between Imc RF lobes. F, Line: Best fit straight line; slope = 0.65, R^2 =0.95. % change in responses directly estimates the strength of inhibition at this OTid neuron due to Imc (here, -35%; section 4.7.4; conventions as in Figure 4.3F). (G-L) Measurement of the strength of net “different lobe” inhibition from Imc → OTid with paired recordings in Imc and OTid. Conventions as in A-F. (G, H) Three-lobed spatial RF of an Imc neuron (left; Figure 4.6E-G; section 4.7.4). RF of OTid neuron overlaps with one of the lobes of Imc neuron’s RF (blue bar). (L) Line: Best fit straight line; slope = 1.18, R^2= 0.9. % change in responses directly estimates the strength of inhibition at this OTid neuron due to Imc (here, 18%; section 4.7.4; conventions as in Figure 4.3F). (M) Population summary of strength of net “gap” inhibition (red; n=12), and strength of net “different lobe” inhibition (blue; n=17) from Imc→ OTid. Open red and blue box plots: reproduced from Figure 4.3L. p<1e-5 (red vs. 0), p=0.35 (blue vs. 0); p=<1e-5 (red vs blue), paired t-tests with HBMC correction. See also Figure 4.6.

We found that Imc neurons exerted strong competitive inhibition onto OTid neurons that encoded locations in the gaps between the Imc RF lobes (Figure 4.5D-F, M-red), but weak competitive inhibition onto OTid neurons that encoded for locations within a different lobe of the Imc's RF (Figure 4.5J-L, M-blue). The mean strength of “gap” inhibition (along elevation) was -55.88% (Figure 4.5M-red, s.d. = 20%, n=12 neuron pairs; p = 1.02 e-6, t-test against 0 with HBMC correction; Figure 4.6IJ-red; section 4.7.4). The mean strength of “different lobe” inhibition was 2.43 % (Figure 4.5M-blue, s.d. = 10.3 %, n=17 neuron pairs; p = 0.35, t-test against 0 with HBMC correction; Figure 4.6IJ-blue; section 4.7.4), and was not distinguishable from the average “self” inhibition (Figure 4.5M- blue vs. open blue; p-value = 0.09; two sample test with HMBC correction). Imc was inactivated effectively in both sets of experiments (median = 98.2% median, 95% CI = [96.46% 99.99%]; Figure 4.6K).

Taken together, these results demonstrated that the Imc implements precisely organized, multi-holed donut-like patterns of inhibition onto the OTid space map, with each Imc neuron's net inhibitory effect creating a pattern complementary to its (multilobed) RF structure (Figure 4.6A).

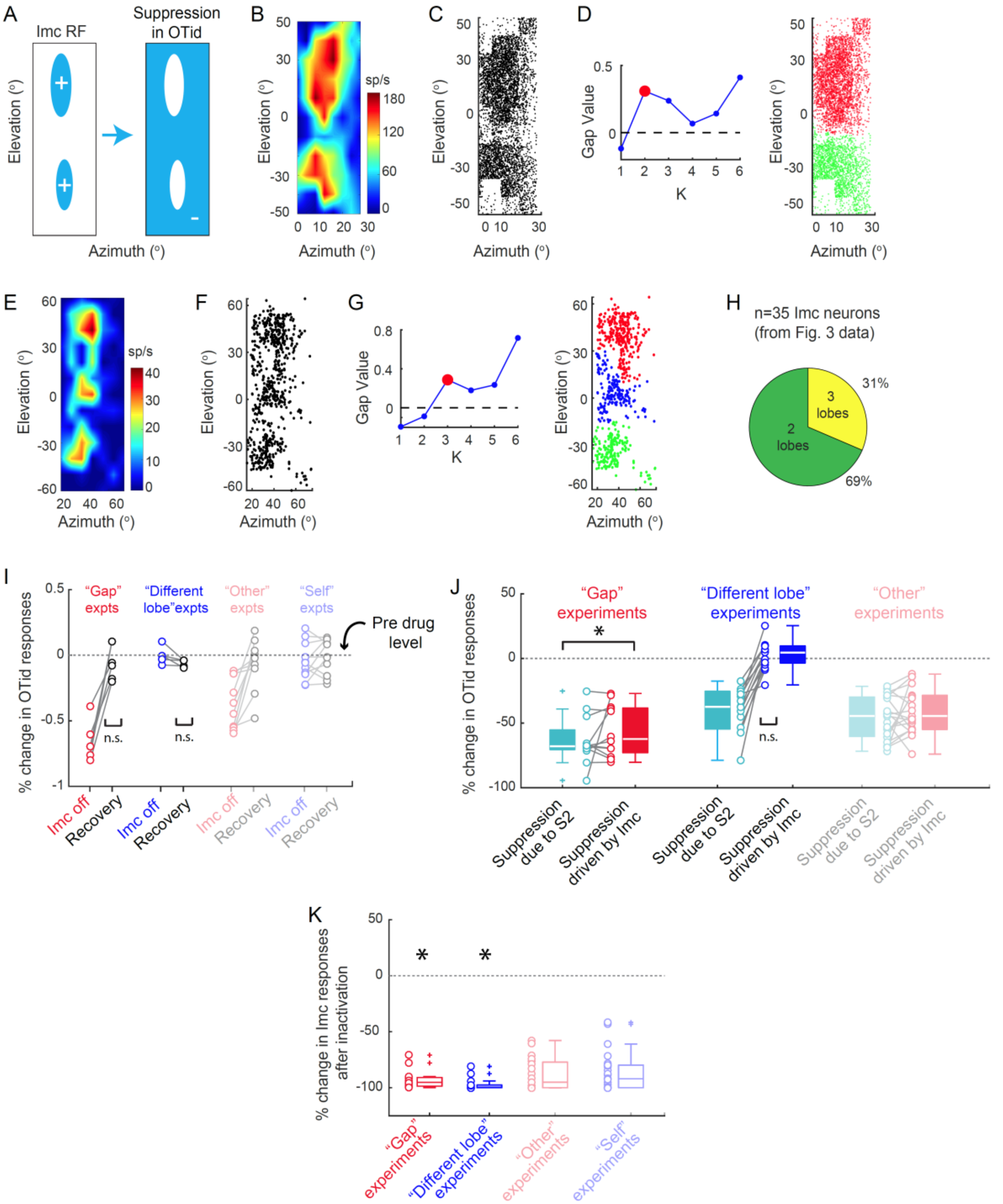


Figure 4.6: Detecting the number of Imc RF lobes, and supporting experimental data for Figure 4.5.

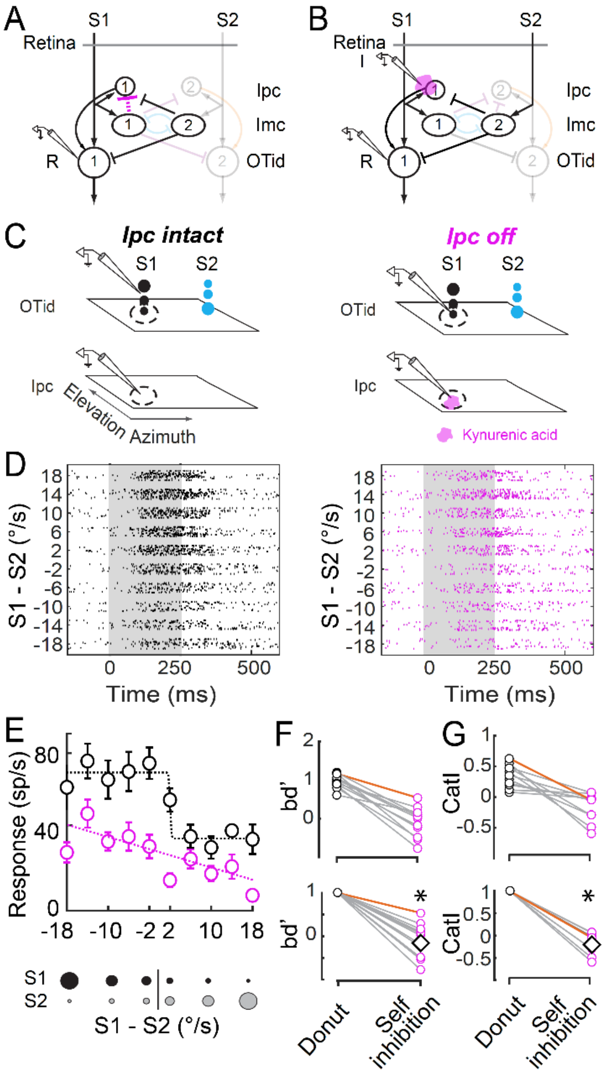
(A) Schematic representation of a multi-holed donut. Left: Two-lobed RF of a putative Imc neuron (left; blue ovals are the RF lobes). Right: Pattern of inhibition created by this neuron in the OTid such that it spares all the locations form which it receives input. This complementary output projection pattern results in two "holes" – areas where no inhibition arrives due to the neuron on the left. **(B)** Imc RF reproduced from Figure 4.5A-left. **(C)** Firing rate map from B converted to density of points in 2-D plane following published procedures [146]. **(D)** Left: Density peaks clustering method [109] is applied to the data in

B, forcing the method to yield either 1 cluster, 2 clusters, … upto 6 clusters. Following that, the gap statistic model selection metric [110] is applied to the clustering results to identify the optimal number of clusters in the data. The optimal number is the number for which the "gap value" exceeds 0 for the first time; here it is 2. Right: Same as B, but with the two colors indicating the two distinct lobes identified by the clustering method (+ gap statistic) as the optimal two best clusters in the RF data. **(E-G)** Same as B-D, but for Imc RF in Figure 4.5G-left; determined to be three-lobed. **(H)** Pie-chart showing proportion of two-lobed vs. three-lobed RFs across all the multi-lobed Imc neurons recorded in this experiments (Figure 4.5). **(I)** Recovery of OTid responses from kynurenic acid iontophoresis for experiments in Figure 4.5. Conventions as in Figure 4.4D; data from Figure 4.4D reproduced here for comparison. Data show recovery ("Gap" expt: p =0.39, t-test of black data against 0; "Other" lobe expt: p=0.29, ranksum test of black data against 0), demonstrating that any effects reported Figure 4.5 are due specifically to drug iontophoresis/Imc inactivation. **(J)** Comparison of suppression provided by Imc with that due to S2 (i.e., the maximum amount of suppression experienced by the OTid neuron in this stimulus protocol) in the "gap" (red data) and "different lobe" (blue data) experiments. Conventions as in Figure 4.4E; data from Figure 4.4E reproduced here for comparison. "Gap" expt: Consistent with [100], nearly all the suppression due to competitor S2 (teal) is supplied by Imc (red) (p= 0.47, teal vs. red, ranksum test). "Different lobe" expt: The Imc neuron that is being inactivated does NOT provide any of the inhibition to the OTid neuron due to S2 (p = 0.35, blue dots against 0, t-test). This clearly demonstrates, consistent with the predictions of [146], that a different Imc neuron exists, which has one RF lobe encoding S2's location, but other RF lobes (if there are others for that neuron) not encoding S1's location, thereby delivering inhibition to that location! This is one of the 'signature' properties for the combinatorial encoding of space described in that study. **(K)** Quantifying the effectiveness of Imc inactivation by kynurenic acid iontophoresis in the "gap" (red) and "different lobe" (blue) experiments. "Gap" expt: p =1.13 e-12 t-test against 0; "Other" lobe expt: p =1.53e-5, signtest. Conventions as in Figure 4.4F; data from Figure 4.4F reproduced here for comparison.

4.5 Donut-like inhibitory motif is required for categorization by barn owl midbrain

Considering the wiring complexity of the (multi-holed) donut-like connectivity between Imc (and Ipc) and OT, we wondered if it served a functional purpose. Specifically, we investigated whether it was necessary for the categorical signaling of the highest priority stimulus by the OTid. To test this, we reasoned that causally introducing "self" inhibition onto an OTid neuron to disrupt the donut-like pattern of inhibition would be most effective if achieved via the indirect, but potent, pathway, i.e., by increasing "self" inhibition onto Ipc, (Figure 4.7A). As the net "self" inhibition at most OTid neurons is weak or zero (Figure 4.3L and Figure 4.5M), activating Imc → Ipc projections (or Imc → OTid projections) for spatially aligned neuron pairs would be infeasible because such projections are unlikely to exist. Instead, we mimicked this effect by focally inactivating Ipc neurons (Figure 4.7B-right),

Figure 4.7: Donut-like inhibitory motif is required for categorical signaling of the strongest stimulus in the barn owl midbrain selection network.



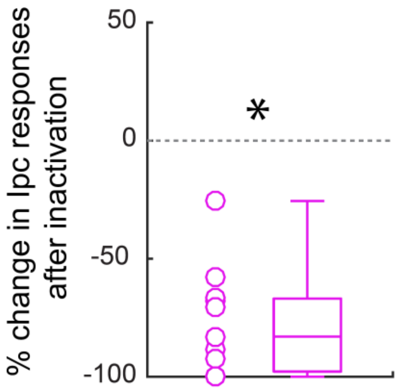
(A) Network model showing the OT-Imc-Ipc circuit in the barn owl. Magenta, dashed arrow: Introduction of "self" inhibition in the Imc→ Ipc pathway; would disrupt donut-like inhibition, but is infeasible (Figure 4.3 and Figure 4.5; text). **(B)** Schematized is iontophoretic inactivation of Ipc neuron aligned with OTid neuron; a feasible causal manipulation that disrupts the donut-like inhibition within channel 1 by de-activating the donut-like motif onto the amplifier. **(C)** Paired recordings in OTid and Ipc such that RF of OTid neuron overlaps with that of Ipc neuron. Conventions as in Figure 4.3C. Stimulus protocol used is the same two-stimulus morphing protocol used in model simulations in Figure 4.1E; the relative strength between S1 and S2 was systematically varied. **(D)** Left: OTid response rasters in the Ipc-intact condition. Right: OTid response rasters in the Ipc-off condition. Distance between OTid and Ipc RF centers = 5°. **(E)** OTid response firing rates, computed from D over the 100-400 ms time window. Black: Ipc-intact condition, magenta: Ipc-off condition. Dashed lines: best fitting sigmoid or straight line to data, chosen based on AIC criterion. Black: AIC (sigmoid) = 29.07, AIC (line) = 40.92; magenta: AIC (sigmoid) = 37.75, AIC (line) = 34.67. **(F, G)** Population summary of effect of Ipc inactivation on bd' (F) and CatI (G); n=11 neuron pairs; data in orange are from example neuron pair in D-E. Top panels: measured values. Bottom panels: Data in top panels replotted after normalizing to Ipc-intact values. Diamond: average of magenta data; $p < 1e-4$ (bd'), t-test, and $p < 1e-4$ (CatI), ranksum tests against 1 with HBMC correction in F and G. See also Figure 4.8.

while recording simultaneously the responses of spatially matched OTid neurons to the two-stimulus morphing protocol used to study categorical signaling in the OTid (Figure 4.7C left vs. right;

protocol shown in Figure 4.7E–bottom, same as Figure 4.1E– bottom; section 4.7.4;[95]).

We found that disrupting the donut-like inhibitory motif (by introducing "self" inhibition onto the Ipc) caused a substantial reduction in categorical signaling in the OTid (Figure 4.7D-E: example neuron pair). Across the population of tested neuron pairs, categorization as well as robust-to-noise signaling were nearly abolished, with the median reduction in *CatI* of 104.7% and a mean reduction in *bd'* of 112.42% (Figure 4.7FG; n=11 neuron pairs; *bd'*: s.d. = 37.8%; p = 3.93e-9, t-test against 1 with HBMC correction; *CatI*: 95% CI of median = [91.52%, 117.92%]; p = 2.55 e-5, ranksum test against 1 with HBMC correction). Note however, that the ability to discriminate stimuli pairs that are far from the selection boundary remains unaffected. Thus, the midbrain spatial selection network not only contains the specialized donut-like inhibitory circuit motif, but also critically depends on it for categorical, robust-to-noise signaling of the highest priority stimulus.

Figure 4.8: Ipc is effectively inactivated.



Quantifying the effectiveness of Ipc inactivation by kynurenic acid iontophoresis in the "necessity of donut-like inhibition" experiment. The inactivation was highly effective: (mean % change in responses = -77.42%, sd =22.82% p =5.3e-7, t-test against 0).

4.6 Divisive normalization model is not effective for generating categorical representations

The structured (multi-holed) donut-like organization of inhibition stands in contrast to another computational mechanism that has been invoked in the decision-making literature, namely divisive normalization [141, 147-149]. It involves the pooling of inhibition due to various active channels, with the inhibitory elements inhibiting each other, and with the pooled inhibition delivered *uniformly* (rather than in a donut-like manner) to the output elements (Figure 4.9A;[147, 148]).

Our results (Figure 4.3, Figure 4.5) show that the midbrain spatial selection network (OT-Imc-Ipc) does not implement pooled divisive normalization for selection across space. To test, more generally, whether the divisive normalization model is effective for generating categorical response profiles, we simulated a circuit model of normalization (Figure 4.9A) and obtained model neuron responses to the two-stimulus morphing protocol (Figure 4.9A; same protocol as in Figure 4.1E and Figure 4.7). *CatI* was substantially lower for this model than that obtained with just the donut-like inhibitory motif (Figure 4.9C right panel: mean *CatI* = -

0.05 for normalization model compared to 0.331 for donut model in Figure 4.1, p = 1.5 e-29, t-test with HMBC correction, gold vs. purple; left panel: p = 1e-12, t-test of gold vs. purple,). Thus, the normalization mechanism is not effective for generating categorical responses (consistent with findings from modeling in visual cortex [150]). Indeed, it was the presence of self-inhibition, specifically, that caused this circuit to be ineffective: a modified version of the circuit in Figure 4.9A which did not include self-inhibition (Figure 4.9B), was very effective, further attesting to the primacy of the donut-like motif for categorization (Figure 4.9C right panel: blue vs. gold data, mean *CatI* = 0.279 for normalization model with donut compared to 0.022 for normalization model, p = 1.15e-14, t-test with HBMC correction; left panel: p = 8.8e-19, t-test of blue vs. gold with HBMC correction). (Incidentally, the donut-like motif is capable also of keeping the responses of the circuit in check, in a sense, 'normalizing' the responses, as the overall drive increases (Figure 4.1E)).

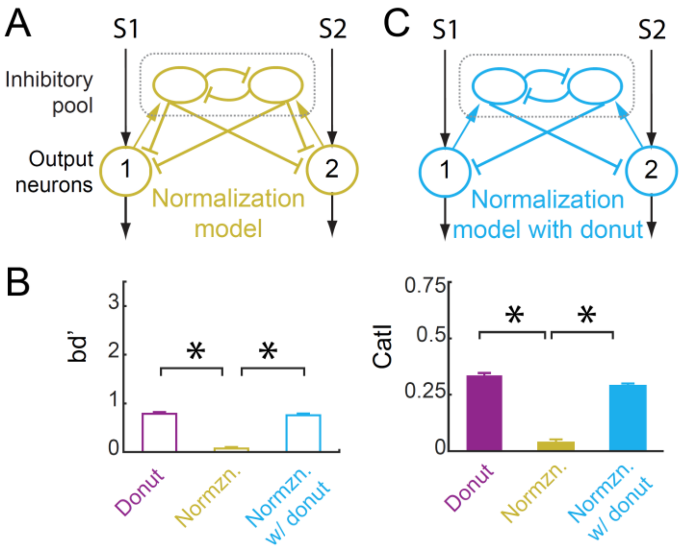


Figure 4.9: Divisive normalization model operating across the space map is not effective for producing categorical responses.
(A) Schematic of normalization circuit model with pooled inhibition and inhibitory feedback [147, 148]. (B) Schematic of normalization circuit sans self-inhibition. (C) Plot of bd' (left panel) and CatI (right panel) computed from the responses of neuron 1 in the normalization model (A; gold data) and the model sans self-inhibition (B; blue data), to the standard two-stimulus morphing protocol (as in Figure 4.1E). For comparison, the bd' and

CatI values for donut-like motif, from Figure 4.1FG, reproduced here for comparison. Right panel (CatI): p = 5.2e-23 (gold vs. purple), p=8.8e-19 (gold vs. blue), paired t-tests with HBMC correction; left panel (bd’): p = 1.9e-31 (gold vs. purple), p=5.13e-31 (gold vs. blue), paired t-tests with HBMC correction.

4.7 Data collection and analysis methods

4.7.1 Data collection and spike sorting

***Neurophysiology*.** For the findings in this chapter, **w**e performed experimental recordings in 7 head-fixed awake adult barn owls viewing a visual screen passively (Tyto alba). The care and housing of animals are as described in section 2.4.1.

Experiments were performed following protocols that have been described previously [75, 100]. Briefly, epoxy-coated, high impedance, tungsten microelectrodes (A-M Systems, 250µm, 5-10MΩ at 1 kHz) were used to record single and multi-units extracellularly in the OTid. Multi-barrel glass electrodes (Kation Scientific, Carbostar– 3LT, 0.4-1.2MΩ at 1kHz) filled with kynurenic acid (a competitive inhibitor of ionotropic glutamate receptors; pH 8.5-9 at a concentration of 40mM) were used to record from and inactivate neurons in the Imc and Ipc. Inactivation was performed using micro iontophoresis by ejecting kynurenic acid with an eject current of -450 nA to -500 nA; data were collected starting 15 min after drug ejection commenced. A retain current of +15 nA was used to prevent leakage of the drug from the tip of the electrodes when drug was not being iontophoresed. Recovery data were measured 15 min after drug ejection was ceased.

Multi-unit spike waveforms were recorded using Tucker Davis Technologies hardware interfaced with MATLAB. The responses of neurons were measured by counting spikes during a time window (typically 100-350 ms) following stimulus onset.

The automated ‘wave-clus’ spike-sorting toolbox was used for spike sorting [151]. We included only those units for analysis for which fewer than 5% of the recorded spikes were within 1.5 ms (inter-spike interval; ISI) of each other.

***OT, Imc and Ipc targeting*.** We navigated to the OT and Imc using methods described in section 2.4.3. The Ipc lies roughly 500-700 um medial to the Imc and its targeting was confirmed based on the neural response characteristics of the neurons (characteristic bursty responses; [104]).

4.7.2 Model details

1. Input output functions

The input output functions (firing rate f, as a function of the saliency, l) of the neurons in the model were simulated using sigmoid functions using previously published methods[139].

$$f(l) = c + s\left(\frac{l^m}{l^m + L_{50}^m}\right)$$

where c, is the baseline firing rate of the neuron; l, is the saliency parameter of the stimulus (e.g. loom speed of the stimulus, loudness of an auditory stimulus, contrast of the stimulus, speed of a moving stimulus) that can vary continuously over a range; s, is the maximum change in the firing rate of the neuron; L_{50} is the saliency value at which the neuron’s firing rate changes by 50% of the maximum change; and m, is a parameter that controls the slope of the sigmoid.

2. Excitatory neurons

The excitatory neurons in Figure 4.1, Figure 4.2, Figure 4.4, and Figure 4.9 were simulated using the following parameters (which are consistent with the parameters obtained by fitting a sigmoid to response functions of OTid neurons [139]:

$$c = 5.3, s = 22.2\ , L_{50} = 11.6, m = \ 2$$

3. Inhibitory neurons

The inhibitory neurons in Figure 4.1, Figure 4.2, Figure 4.4, and Figure 4.9 were simulated using the following parameters (which are consistent with the parameters obtained by fitting a sigmoid to response functions of Imc neurons [139]:

$$c = 5, s = 15\ , L_{50} = 8, m = 10$$

4. Recurrent excitation neurons

The excitatory neurons that provide recurrent amplification in Figure 4.4 were simulated using the following parameters (which are consistent with the parameters obtained by fitting a sigmoid to response functions of Ipc neurons; [152]):

$$c = 8.4, s = 36\ , L_{50} = 5.8, m = \ 3.3$$

The inhibition sent from the inhibitory neurons onto excitatory neurons is modeled using input and output divisive factors as below using previously published methods [139].

$$f = \left(\frac{1}{s_{out}} + 1\right).\left(\frac{c}{s_{in}+1} + s\left(\frac{l^m}{l^m + L_{50}^m + s_{in}^m}\right)\right), \text{ where}$$

- $s_{in} = d_{in}.I, s_{out} = d_{out}.I$ are the input and output divisive factors.
- $d_{in}\ and\ d_{out}$ are parameters that control the strength of input and output division and I is the output (firing rate) of the inhibitory neuron sending the inhibition.

The value of these parameters chosen were $d_{in} = 0 \, and \, d_{out} = 0.25$ consistent with previously published methods (see Figure. 5D in [139]).

5. Feedback inhibition

In models in which the inhibitory neurons inhibit each other (e.g. Figure 4.1C, bottom left panel, green connections; feedback model), the feedback inhibition was modeled as below using previously published methods[139].

$$I(t) = \left(\frac{1}{i_{out}} + 1\right).(\frac{c}{i_{in} + 1} + s\left(\frac{l^m}{l^m + L_{50}^m + i_{in}^m}\right))$$

$$i_{in}(t) = r_{in}.I'(t-1), \quad i_{out} = r_{out}.I'(t-1)$$

where I' is the output (firing rate) of the other inhibitory neuron at time $(t-1)$.

These equations were iteratively applied until there was no further change in the output of the neurons (i.e., steady state was reached).

The values of the feedback parameters used were $r_{in} = 0.8, r_{out} = 0.01$ consistent with previously published methods. We also varied these two parameters (varying feedback; Figure 4.2A) to study their effect on boundary discriminability and categorization index.

6. "Self" inhibition and Donut.

In models in which the excitatory neuron receives inhibition from more than one inhibitory neuron (e.g. Figure 4.1C, top left panel; Baseline model), the inhibition from these sources was combined as below.

$$f = \left(\frac{1}{s_{out1}} + 1\right).\left(\frac{1}{s_{out2}} + 1\right).(\frac{c}{s_{in1} + s_{in2} + 1} + s\left(\frac{l^m}{l^m + L_{50}^m + s_{in1}^m + s_{in2}^m}\right))$$

where, $s_{in1} = d_{in}.I_1, s_{out1} = d_{out}.I_1, s_{in2} = (s * d_{in}).I_2, s_{out2} = (s * d_{out}).I_2$

I_1 is the output (firing rate) of the inhibitory neuron 1 and I_2 is the output (firing rate) of the inhibitory neuron 2.

The parameter value s controls the strength of self-inhibition, and ranges between 0 and 1. In models that have 'donut-like connectivity', s is set to 0 (Figure 4.1C, top right panel; Donut model). In models which do not have the donut-like connectivity (Figure 4.1C, top left panel; Baseline model), s is set to 1.

We also vary the value of s systematically between 0 (donut) and 1 (maximum self-inhibition) to study the effect of the strength of "self" inhibition on the boundary discriminability and categorization index (Figure 4.2B).

*7. **Recurrent excitation***

In the models with recurrent excitation, (Figure 4.1D), the output of the neuron is scaled by a factor (k, k= 2.5 in Figure 4.1FG). We also vary the scaling factor to study its effect on boundary discriminability and categorization index (Figure 4.2C).

In the model in Figure 4.4, the output of the neuron which receives recurrent amplification is modeled as below.

$$f(l) = (1 + e_{out}).\left(\frac{1}{s_{out} + 1}\right).\left(\frac{c}{s_{in} + 1} + s\left(\frac{l^m}{l^m + L_{50}^m + s_{in}^m}\right)\right)$$

where, $e_{out} = e_a.A,\quad e_a = 0.01$, and A is the output of the neuron sending the amplification.

The value of the parameter is chosen as 0.01 to yield results consistent with the amplification effect of Ipc on OT responses as reported in previously published work (Ipc inactivation results in a 31% decrease in the OTid responses on average; [142]).

8. Normalization model

To implement the model in Figure 4.9 (normalization model), we used the input-output functions and the effect of divisive inhibition described above. The primary feature of these models that is different from models in Figure 4.1, Figure 4.2 and Figure 4.4 is that the excitatory neurons send inputs to inhibitory neurons, which then inhibit those excitatory neurons. The output of an excitatory neuron at time t is calculated by applying divisive inhibition (from the inhibitory neurons at time t-1) to its activity at time t-1. This activity is then used to calculate the activity of the inhibitory neurons at time t as below.

$$I(t) = (1 + e_{out}).I(t-1)$$

where, $e_{out} = e_a.A, \quad e_a = 0.01$, and A is the output of the neuron sending the input at time t-1.

This process is repeated iteratively until steady state activity is obtained. The initial activities (at t=0) of the inhibitory neurons are set to their baseline level, and of the excitatory neurons are calculated from the stimulus inputs without any divisive inhibition.

4.7.3 Calculating boundary discriminability and categorization index

For each of the models in Figure 4.1, Figure 4.2 and Figure 4.4, we used a two-stimulus strength morphing protocol described in previously published work [95]. We presented stimulus S1 at location 1 and stimulus S2 at location 2. As the strength of the first stimulus was increased, the strength of the second stimulus was decreased (Figure 4.1E, Figure 4.7E). Responses of the model output neuron (#1) were simulated using this protocol. Random noise from a standard normal distribution was added to the responses. The variance of the noise added depended on the mean value of the responses (m) and a fano factor value (ff): variance = ff *m. This was repeated 30 times (to mimic 30 reps of data collection from a 'neuron') and was used to compute the response profiles (mean +/- s.e.m.) of that neuron from different model circuits. Similarly, response profiles for 50 neurons were obtained for each model circuit. For all the model runs reported in Figure 4.1, Figure 4.2, Figure 4.4 and Figure 4.9, we used a fano factor value of 6. We also varied the fano factor value to test the effect of noise on model performance (Figure 4.2D).

This stimulus presentation protocol results in 2 categories (Category 1: S1 > S2 and Category 2: S1 < S2). We measured the boundary discriminability (bd') of the responses of neuron 1 by calculating the d-prime between the responses of neuron to stimuli pair straddling either side of the selection boundary and at a distance of 3 units from the boundary as:

$$bd' = \frac{\mu_1 - \mu_2}{\sqrt{0.5\left(s_1^2 + s_2^2\right)}}$$

where, μ_1 and s_1 are the mean and the standard deviation of the responses of the neuron to the stimuli pair from category 1 near the boundary; and μ_2 and s_2 are the mean and the standard deviation of the responses of the neuron to the stimuli pair from category 2 near the boundary.

To compute the categorization index, we compared two quantities: (modified from previously published work; [95, 123]) (a) the mean within-category d-prime (WCD') between the responses of the neuron to pairs of stimulus-pairs (S1 and S2) that are in the same category, and (b) the mean between-category d-prime (BCD') between the responses of the neuron to pairs of stimulus-pairs that are in different categories. The pairs are chosen while ensuring that

(i) the number of pairs used to calculate both these metrics are the same, and (ii) the distribution of the distances between the chosen pairs for calculating both the metrics are the same. The categorization index is calculated from these two metrics as:

$$Categorization\ index\ (CatI)\ = \frac{mean(BCD') - mean(WCD')}{mean(BCD') + mean(WCD')}$$

CatI= 1 indicates idealized, step-like responses; = 0 indicates linear, non-categorical responses; < 0 indicates better discriminability within category than between categories.

4.7.4 Data collection protocol

Visual stimuli used here have been described previously [75, 95]. Briefly, looming visual dots are flashed at different locations on a tangent TV monitor in front of the owl. Looming stimuli were dots that expanded linearly in size over time, starting from a size of 0.6° in radius. The speed of the loom was decided based on the stimulus protocol and varied between 9.6°/s and 21.6°/s. Visual stimuli were presented for a duration of 250ms with an inter stimulus interval of 1.25s to 4s.

1. Receptive fields (RFs)

For measuring spatial RFs of Imc, OTid and Ipc neurons, a single stimulus of a fixed contrast was presented at the sampled locations. The order of locations at which the stimulus was presented was randomized to minimize adaptation. A location is considered to be inside the RF if it evokes responses significantly different from that of baseline response. All other locations are considered to be outside the RF.

The following stimulus protocols were used for measuring data reported in Figure 4.4 and

Figure 4.6.

2. ***Paired OTid and Imc data collection***

i. *"Other" inhibition*

To measure the strength of "other" inhibition sent from an Imc neuron to a distant location in the OTid space map, we simultaneously recorded from the Imc neuron (with a multi-barrel glass electrode filled with the kynurenic acid), and a spatially misaligned "other" OTid neuron (with a tungsten electrode) as described below; we ensured that the half-max of the RF of the OTid neuron lay outside (did not overlap with) the half-max of the Imc RF.

For measuring "other" inhibition, we recorded the following data curves together in an interleaved manner.

a) Tuning curve (TC) centered at the OTid RF peak: A 1-dimensional spatial tuning curve centered at the peak of the OTid RF. The stimulus had a loom strength of 9.6°/s.

b) Tuning curve centered at the OTid RF peak + competing stimulus centered at the Imc RF peak (TCC): The same curve as in a), but along with a (distant) competitor positioned at the peak of the Imc RF. The strength of the competing stimulus was chosen to be 19.2°/s. This second competing stimulus drives the Imc neuron, which sends strong competitive inhibition to the OTid neuron. Since the competing stimulus was more salient than the stimulus driving the OTid neuron, the responses of the OTid neuron were strongly suppressed consistent with previous published results [75].

c) Tuning curve (TC) centered at the Imc RF peak: 1-dimensional spatial tuning curve centered at the Imc RF. The loom strength of the stimulus was chosen to be 9.6°/s.

We measured the above 3 curves both when the Imc neuron is intact ('baseline' condition) and focally inactivated ('inactivation' condition) and compare the responses as below. In a subset of the data, we also measured the curves after the responses of the Imc neurons recovered (Figure 4.4D, 'recovery' condition) from focal iontophoretic inactivation. Inactivation curves were measured 15 min after the drug ejection was started. Recovery curves were measured 15 min after the drug ejection was stopped.

We analyzed the data from these three curves as below.

First, a distant competing stimulus is known to typically suppress OTid responses [75, 107], something that we confirmed by comparing the responses of OTid neurons to the TC (curve a) and TCC (curve b); Figure 4.4E. Notably, it is also known that some OTid neurons do not show such suppression, and any such neurons were excluded from our analyses [94].

Next, to measure the amount of Imc inactivation, we compared the responses of the Imc neurons to the TC measured at the Imc RF peak (curve c) in the baseline condition versus the inactivation condition. Kynurenic acid was able to effectively shut down Imc responses: (Figure 4.3F; red, median strength of Imc inactivation = 95%, 95% CI = [87%, 103%], p = 3.8e-6, sign test, n=19).

Finally, to measure the strength of "other" inhibition, we compared the TCC responses (curve b) in the baseline condition and the inactivation condition. In the baseline condition, the Imc neuron was intact, driven by the competing stimulus and sent strong inhibition to the OTid neuron. In the inactivation condition, the Imc neuron was shut silenced, as a result of which the OTid neuron was released from inhibition and exhibited an increase in responses.

Any observed increase in OTid responses quantified the amount of suppression due to Imc onto the OTid location, thereby directly estimating the strength of net "other" inhibition: % change in OTid responses = 100* (responses in Imc intact condition – responses in Imc off condition)/responses in Imc off condition. To obtain an accurate estimate of the %

change, we considered the responses to stimulus S1 (in the two conditions) at all locations inside the RF as follows. We fit a straight line to the plot of OTid responses to S1 in the Imc intact vs. Imc of conditions (Figure 4.3F), calculated the slope of the best-fit line, and used it to compute the average value of % change as: %change in OTid responses = 100 * (slope-1). This procedure was also used to quantify "self" inhibition, "gap" inhibition and "different lobe" inhibition (below).

ii. *"Self" inhibition*:

To measure the strength of "self" inhibition sent from an Imc neuron to a matched location in the OTid space map, we simultaneously recorded from the Imc neuron (with a multi-barrel glass electrode filled with the kynurenic acid), and a spatially aligned "self" OTid neuron (with a tungsten electrode) as described below; we ensured that the half-max of the RF of the OTid neuron overlapped with the half-max of the Imc RF.

We recorded a spatial tuning curve centered at the peak of the OTid (and therefore, Imc) RF. The loom speed of the stimulus used was 9.6°/s; this stimulus drives both the OTid and the Imc neuron.

We compared the responses of the OTid neuron in the Imc-intact (baseline) and Imc inactivated condition and the difference quantified the strength of "self" inhibition (Figure 4.3L, blue).

To measure the amount of Imc inactivation, we compared the responses of the Imc neuron in the baseline and the inactivation condition as before. Kynurenic acid was able to effectively shut down Imc responses: (median strength of Imc inactivation = 92%, 95% CI = [86% 98%], p = 7.5e-9, sign test).

iii. *"Gap" inhibition*:

To measure the strength of "self" inhibition sent from an Imc neuron with a multilobed RF to the locations in the OTid space map between the RF lobes (gaps), we simultaneously recorded from the Imc neuron (with a multi-barrel glass electrode filled with the kynurenic acid), and "gap" OTid neuron (with a tungsten electrode), the RF (half-max) of which was located in the gap between the half-max of the lobes of the Imc RF (Figure 4.5A-C). We recorded the same set of curves as in the "other" inhibition case (see above) from the gap OTid neuron, and performed similar analyses on the data to quantify the strength of "gap" inhibition.

iv. *"Different lobe" inhibition*:

To measure the strength of inhibition sent "from" locations within one lobe of an Imc neuron with a multilobed RF to the locations in the OTid space map within other lobes of the Imc RF ("different lobe" inhibition), we simultaneously recorded from a multilobe Imc neuron (with a multi-barrel glass electrode filled with the kynurenic acid), and an OTid neuron (with a tungsten electrode), the RF (half-max) of which overlapped with one of the lobes of the Imc RF (Figure 4.5G-I). Then we measured the same set of curves as in the "other" case while ensuring one additional detail. We recorded the TC curve (curve a) with stimulus (S1) centered at the OTid RF peak (and therefore within one of the lobes of the Imc RF as well). For the TCC (curve b), the second stimulus S2 was centered at the peak of a different lobe of the Imc neuron's multilobe RF. Thus, both stimuli excited the Imc neuron (during curve b) as they lay within its RF, but only S1 excited the OTid neuron and S2 served as a (distant) competitor from the its perspective.

An OTid neuron in this configuration was defined as a “different lobe” OTid neuron; the half max of its RF overlapped with the half max of one of the lobes of the Imc RF, but not of other lobes. (Note that by definition, every “different lobe” OTid neuron is also a potentially “self” OTid neuron, and admits to the use of the appropriate stimulus protocol to measure “self” inhibition. However, the opposite is not true, since “self” OTid neurons can be identified for Imc neurons with single lobed RFs as well, but “different lobe” OTid neurons are only defined for multilobe Imc neurons).

We applied similar analyses as in the “other” case to the data from “different lobe” OTid neurons and quantified the strength of “different lobe” inhibition.

3. Paired OTid and Ipc data collection

To test if the donut-like inhibitory motif is required for robustness-to-noise and categorization, we made paired recordings at spatially aligned Ipc and OTid neurons. We used a strength morphing stimulus protocol described in previously published work [95]. Briefly, we presented one stimulus (S1) inside the RF of the OTid neuron (and also Ipc neuron because their RFs overlap). Simultaneously, we presented a competing stimulus (S2) 30° away along azimuth from S1. As the strength of the S1 decreased, the strength of S2 stimulus increased (Figure 4.7E).

We applied the same analyses described above for the model in Figure 4.1 and Figure 4.2 to compute the boundary discriminability and the categorization index for experimental data reported in Figure 4.7.

4.8 Discussion

Our results discover the power of a donut-like inhibitory motif for generating categorical representations and its ability to convert even linear response profiles to categorical ones (Figure 4.1E-inset: orange/brown vs. purple). This compact motif does not rely on mechanisms of plasticity for its operation; it is able to dynamically generate categorical responses from continuously varying inputs, on the fly. Such dynamic categorization is critical in various behaviors including selection for spatial attention, value-based decision-making, action selection, etc. Here, the donut-like inhibitory motif generates categorical responses by operating on a representational substrate in which individual stimuli are encoded with neural activity that is proportional to their net priority [59, 79, 144]. It readily generalizes to other instances of selection in which stimulus options are encoded with neural activity that is proportional to their net attractiveness or importance: for instance, subjective-value of an option in the case of value-based decision-making [153-155], degree of membership of a stimulus in the case of perceptual categorization [125, 156], etc. Additionally, it could also play a functional role in sculpting categorical responses that arise through learning[123, 128, 157, 158]. Taken together, the donut-like inhibitory motif may be a neural circuit module common to various forms of categorical selection and decision-making.

It is important to note that categorical neural representations are not inconsistent with selection behavior that exhibits smooth, sigmoidal psychometric curves, as is typically found to be the case in selection and decision-making tasks [93, 155, 159, 160]. This is in line with previous findings that neurons can encode information more effectively than the animal as a whole, with behavior being a result of (noisy) aggregation of activity across neurons [161]. In the context of selection, then, we would predict that categorical neural representations are necessary for the observed psychometric curves in decision-making tasks, with their disruption causing psychometric curves to become shallower, and causing performance to worsen specifically around the selection boundary.

Here, we described the donut-like motif in the context of selection across space, which is underpinned by well-organized topographic maps of space. However, for many forms of selection, such organized maps do not exist, with olfactory categorization being an extreme example [125], involving dense combinatorial coding of odors. The findings from this study suggest that detailed investigations of the properties and connectivity of inhibitory neurons in those cases may be a fruitful endeavor. In the subcortical circuit studied here, which is conserved across all vertebrates [79, 85], the long-range suppression that is necessary for the donut-like motif is implemented by inhibitory neurons with far reaching projections. Alternatively, it could also be implemented, for instance, in cortical circuits, through long-range excitation contacting local inhibitory neurons. Additionally, our results testing different implementations of the donut-like motif to match leading classes of models of decision-making support the generality of this neural circuit motif for implementing categorization across brain areas.

A seeming alternate mechanism for categorization that we did not examine in detail is the biophysical specification of input-output functions of the neurons to be step-like. However, the limited power of this mechanism for producing categorical responses in general circumstances, together with the past demonstration that the steepness of the i/o functions of neurons can be uncorrelated with whether selection is categorical [94], made this alternative not compelling. Another alternative are highly recurrent networks, which have been shown in modeling to be capable of generating categorical outputs from multiplexed encoding through dynamics [162, 163]. However, because it is difficult to extract specific, testable neural circuit mechanisms from such a network, and in light of the recently reported counterpoint to mixed-selectivity descriptions [126], we focused, here, on structured circuit mechanisms. We considered explicitly the alternatives of feedback inhibition and recurrent amplification. Contrary to prior proposals [137, 139], feedback inhibition, by itself, was not effective for producing categorical response profiles, nor is recurrent amplification by itself, nor the two of them combined. However, if present along with the donut-like motif, they enhanced its efficacy for categorization. Our results, therefore, illuminate the potential mechanistic reason for the success of various past models that have accounted for categorical responses, namely, their inclusion of donut-like inhibition (rather than of amplification or feedback inhibition). Considering that other important functions have been proposed for these two motifs - the implementation of a flexible selection boundary (feedback inhibition; [129, 139, 164]), and evidence accumulation (recurrent amplification; [131, 137]), selection circuits may need to include these motifs. Together with the donut-like inhibitory motif, they can then effectively implement attractor dynamics for flexibly categorical decision-making [125, 132, 137, 139]. An intriguing open question in this context, even in the avian midbrain selection network, is how the wiring of the exquisitely organized (multi-holed) donut-like connectivity is achieved in the brain.

5. Conclusions and Future work

5.1 Summary and general discussion

The goal of this dissertation was to uncover the neural circuit mechanisms of some of the essential computations underlying stimulus selection. Using extracellular recordings, causal manipulations (with drug iontophoresis) and computational modeling, we uncovered the mechanisms of two computations that are essential for stimulus selection (location-invariance and categorization).

1. In chapter 2, we presented the discovery of unusual multilobed RFs in the Imc. This is the first report of such multilobed encoding in early visual sensory brain areas to the best of our knowledge. We systematically characterized the properties of multilobed RFs and the topographic nature of encoding of space in the Imc. We found that whereas the azimuthal space in encoded topographically along the rostrocaudal axis of the Imc, there appears to be no systematic organization of encoding locations along the elevation. Lobes of RFs of Imc neurons appear to be randomly distributed along

elevation a) within individual Imc RFs, b) across RFs of neurons measured at the same recording site, and c) across RFs of neurons measured along the dorsoventral axis of the Imc. Based on theoretical calculations, we predicted that in order to implement selection at all possible pairs of stimulus locations, multilobed encoding would be a necessity in the face of scarcity in the number of neurons available for encoding space. We then tested experimentally if this prediction held true in the barn owl Imc, and found that it did: Imc had a scarcity in the number of neurons available for encoding elevation space, but sufficient number of neurons to encode azimuthal space (consistent with multilobed encoding in elevation but not azimuth).

2. In chapter 3, we investigated the mechanisms used by scarce Imc multilobe Imc neurons to solve stimulus selection at all possible pairs of stimulus locations, a challenging computational problem. We set this up as an optimization problem and discovered that although distributed seemingly randomly, the locations of the Imc RF lobes are exquisitely optimized across the high-firing Imc neurons, allowing Imc to solve selection at all location-pairs using a combinatorial neural inhibition strategy. With experiments, we demonstrated that the owl Imc satisfies all the key properties predicted by the computational model for implementing combinatorial neural inhibition for stimulus selection across space. Further, we showed that this strategy minimizes metabolic and circuit wiring costs, a principle that also accounts for observed asymmetries between azimuthal and elevational coding.
3. In chapter 4, we investigated the neural circuit mechanisms of generating categorical representations. Categorical representations are ubiquitously found across brain areas and form the basis of various behaviors such as selection for spatial attention, decision-making, action selection etc. However, the neural circuit mechanisms controlling such

representations are obtained are relatively less understood. We showed that a 'donut-like' inhibitory mechanism, in which each competing option suppresses all options except itself, is highly effective at generating categorical responses. It far surpasses motifs of feedback inhibition, recurrent excitation, and divisive normalization invoked frequently in decision-making models. We demonstrated experimentally not only that this mechanism operates in the midbrain spatial selection network in barn owls, but also that it is required for categorical signaling by it. Additionally, the functional pattern of neural inhibition in the midbrain forms an exquisitely structured 'multi-holed' donut consistent with this network's combinatorial inhibitory function reported in chapters 2 and 3.

5.1.1 Generality of combinatorially optimized coding beyond the owl and the Imc

Conceptually, combinatorial patterns of inhibition are a consequence of the RF lobes of model Imc neurons being widely distributed and arbitrarily spaced in the optimal solutions (Figure 5.1BD): restricting RF lobes to only nearby locations substantially limits the number of available RF configurations, potentially precluding optimal solutions. Stated equivalently, the combinatorial *inhibition* strategy arises because of the combinatorial *coding of space* by the model Imc neurons: Individual neurons do not always encode only for neighboring locations (Figure 5.1BD), and conversely, nearby locations are not always encoded by 'nearby' neurons (no matter what permutation of neurons is considered; Figure 3.4B and Figure 5.1BD). Therefore, when two stimuli are presented, two groups of the inhibitory Imc neurons are activated in a non-ordered fashion, resulting in a combinatorial pattern of inhibition. This combinatorial coding of space is conceptually distinct from the major population neural coding schemes described thus far in the literature.

For instance, in population vector coding, multiple neurons with overlapping, single-lobed tuning curves (or RFs) are activated to encode feature values such as stimulus locations, motion direction, etc., with high precision [116-119]. With such coding, it is always possible to order these RFs along the feature axis such that neighboring values of features are always encoded by functionally 'local' subsets of neurons (Figure 5.1AC). In contrast, neurons with multilobed RFs cannot be ordered this way: some neurons always code also for distant locations (Figure 5.1BD), and selection for a given location-pair cannot be guaranteed to be solved by only a 'local' subset of neurons (Figure 5.1BD).

A population coding scheme reported in the literature that does involve multilobed encoding as well as the activation of non-local neural subsets is the combinatorial coding of odors by olfactory receptor neurons [120]. However, whereas assorted and extensively intersecting subsets of neurons are activated to encode odors, no inherent constraints that necessitate the optimization of the relative positioning of these RF lobes across neurons have been reported. In contrast, in the combinatorially optimized coding reported here, the placement of RF lobes needs to be optimized across neurons, and is exemplified by the lobe-overlap property (Figure 3.2E).

For this same reason, our scheme also differs from the encoding of space by entorhinal grid cells: the firing fields of different grid cells are not inherently yoked to one another [165, 166]. In addition, each grid cell has a large number of highly organized firing fields, unlike the few, and arbitrarily placed, RF lobes of Imc neurons.

Finally, combinatorially optimized coding also stands in direct contrast to the sparse, orthogonal coding by an overcomplete set of neurons reported in many brain areas [167, 168]. Imc's coding, instead, involves promiscuous, non-orthogonal representation of space by an under-complete set of neurons.

The problem of stimulus selection at all location-pairs with limited neurons, which yields combinatorially optimized coding in Imc, belongs to the same (np-complete) class of computationally complex problems as the traveling salesman problem and the minimum spanning tree problem [169, 170]. Although the brain solves this problem naturally, exactly how Imc's optimized, multilobed RFs are specified during neural development is an intriguing open question and a subject for future work.

The discoveries, here, of multilobed visual representation, combinatorially optimized population coding, and an efficient inhibitory solution for a critical brain function (namely, stimulus selection) have come from the systematic study of the functional response properties of inhibitory neurons in the owl Imc, an area proposed as a critical processing critical processing hub for stimulus selection for attention [87, 99, 100].

Although Imc is conserved across the midbrain of all vertebrates [85, 98, 171], the functional properties of this nucleus of emerging importance have not been studied in any vertebrate other than the barn owl thus far. The biological advantages of the unusual coding strategy elucidated here suggest that combinatorially optimized coding by sparse inhibitory neurons may be a solution employed generally by the vertebrate midbrain to achieve spatial selection across all locations.

Additionally, this computational strategy extends naturally to any selection problem in which the choice must be made no matter what the specific values are of other stimulus features such as orientation, color, odor identity, etc. One example is feature based pop-out [172] in which, for instance, a bar-shaped visual stimulus of a particular orientation stands out (is "selected" neurally and behaviorally) amongst a background of several bars of a different orientation. Since pop-out must and does operate effectively *no matter* what the absolute orientations of the bars are (as long as their relative orientations are distinct), the underlying circuit mechanism must have the ability to compare responses of neurons encoding for different, and in fact - all possible, pairs of orientations, much like the Imc helps compare responses of OT neurons encoding for different (all possible pairs of) locations.

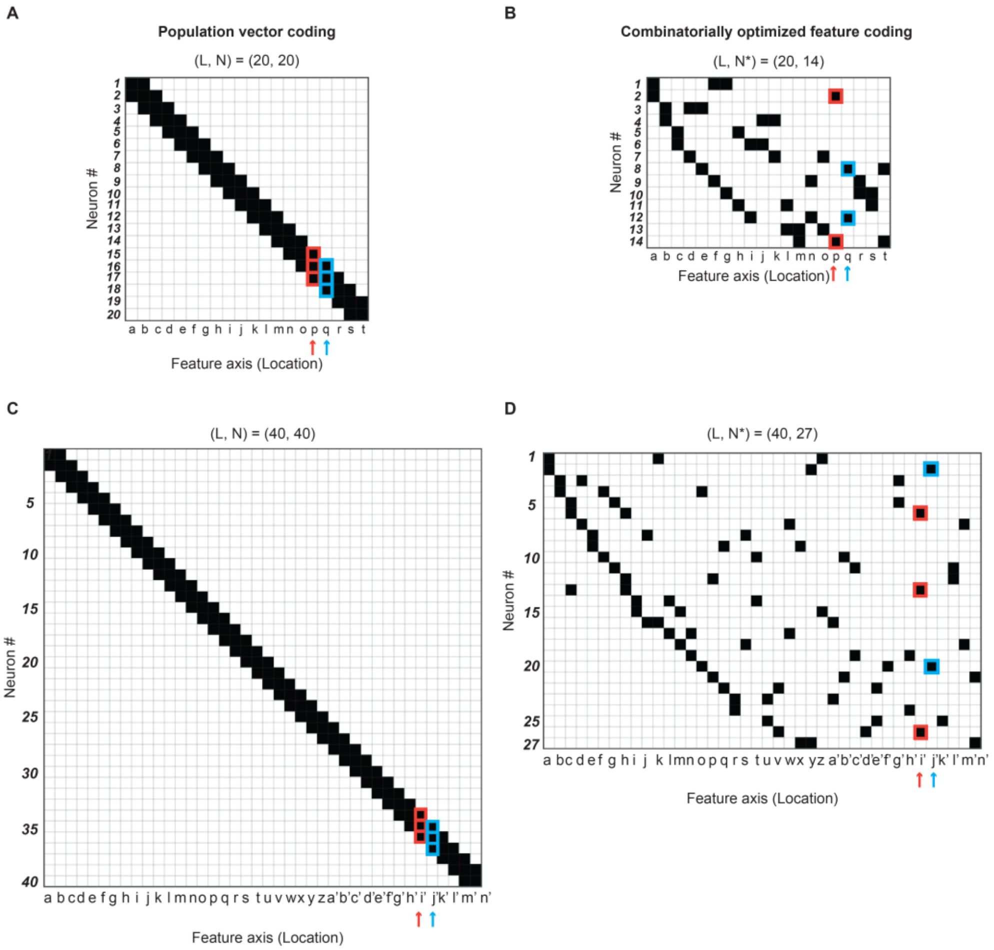


Figure 5.1: Illustration of conceptual differences between traditional population vector coding versus combinatorially optimized feature coding of space.

(A) Population vector coding. Schematic illustration (heat map) of the RFs of 20 single-lobed neurons with overlapping RFs, encoding 20 feature values (say, locations). Neurons are numbered from 1 to 20 (rows), locations are denoted by alphabets (a to t; columns). Black indicates the locations at which each neuron is active. The RF of a given neuron (row) can be read out by looking at the black pixels along that row; the neurons activated by a stimulus at a particular location (column) can be read out by looking at the black pixels along that column. It is evident that each stimulus at a particular location (column) is encoded by a functionally 'local' group of neurons from the ordered set. For instance, stimulus at location p (indicated by red arrow) is encoded by neurons 15, 16, and 17 (indicated in red). Stated equivalently, each neuron encodes only for nearby locations (for instance, row #8). In addition, stimuli at 'nearby' locations (for instance, p and q, indicated by red and blue arrows, respectively) are always encoded by 'nearby' neurons (15,16,17; and 16,17,18, respectively). These features hold true in this canonical ordering of the neurons, in which the RFs cover systematically locations (feature-values) from one end to the other in space (feature space), and this canonical ordering is always possible for such single-lobed RFs. (B) Combinatorially optimized feature coding. Schematic illustration of an optimal model solution (i.e., RFs of inhibitory neurons with optimized overlap) that solve selection across L = 20 locations with N* = 14 neurons. Conventions as in (A). It is evident that not every stimulus location (column) can be encoded by a functionally 'local' group of neurons from the set. For instance, stimulus at location p (indicated by red arrow) is encoded by distant neurons 2 and 14 (indicated in red). Stated equivalently, each neuron does not only encode for nearby locations, rather, it can encode for arbitrarily distant ones (for instance, row #8). In addition, 'nearby' locations cannot always be encoded by 'nearby' neurons. For instance, p and q, indicated by red and blue arrows, respectively, are encoded by widely distributed neurons across the population (2,14; and 8,12, respectively). These features are illustrated here for one particular ordering of the neurons: in order to facilitate comparison with (A), neurons have been numbered such that lower numbers correspond to neurons for which at least one RF lobe occurs earlier than the RF lobes of neurons with a higher number. However, these features hold true no matter the ordering of the neurons, in other words, there is no ordering of the neurons such that 'nearby' feature values are encoded only by 'nearby' neurons. (Note: in (A), the maximum number of pixels in a neuron's RF was chosen to be 3, to match the maximum number of lobes in the RFs of multilobe neurons in (B).) (C) Another illustration of population vector coding using overlapping single-lobed RFs, but with 40 locations and 40 neurons; conventions as in (A). (D) Another illustration of Combinatorially optimized feature coding with an optimal model solution using overlapping multilobed RFs, but with 40 locations and 27 neurons (N* = 27 neurons for L = 40 locations). Conventions as in (B).

Our results suggest that a careful examination of the encoding properties of inhibitory neurons in cortical as well as sub-cortical areas may reveal combinatorially optimized coding as a widespread strategy in the brain for efficient, feature-invariant stimulus selection and decision-making under metabolic and anatomic constraints.

5.1.2 Generalizable implementation of the donut-like motif

Both in the barn owl brain and in the various models considered in Figure 4.1, Figure 4.2, Figure 4.4, and Figure 4.9, the circuit architectures included feedforward inhibition from inhibitory neurons to the 'output' neurons, with the donut-like motif instantiated as the absence of feedforward self-inhibition. Several models of decision-making in humans and monkeys do include a feedforward inhibitory component [173-176]. Examining those models in light of our findings reveals that they indeed implement a donut-like motif, although this aspect of the circuit is not highlighted in the models.

However, another established class of models of decision-making does not include feedforward inhibition, but only involves inhibition in a recurrent path between the competing options [132, 177]. To test if feedforward implementation of the donut-like motif was necessary for categorization or whether its efficacy generalized to recurrent implementations as well, we simulated a version of our circuit model that did not include feedforward inhibition, and that implemented the donut-like motif only via a recurrent route (Figure 5.2AB). For completeness, we simulated this model both without (Figure 5.2A) and with amplification (Figure 5.2B). To implement this model, we used the same methods as for Figure 4.9. For the model with recurrent amplification (Figure 5.2B), the output of the neuron is scaled by a factor (k, k= 2.5). We found that this implementation of the donut-like motif also successfully produced categorical representations (Figure 5.2C; CatI=0.29 (black), 0.34 (light blue)). Nonetheless, the feedforward implementation of the donut-like motif (Figure 4.1CD bottom right panels) offered additional benefits to categorization (Figure 5.2C, right panel: black (Figure 5.2A) vs. dark blue (Figure 4.1C-bottom–right), p = 1.7e-25; light-blue (Figure 5.2B) vs. orange (Figure 4.1D-bottom–right), p = 2.5e-31; paired t-tests with HBMC correction), indicating that if a neural circuit (for instance, the avian midbrain selection network) is able to implement this version of the donut, it can do better than one with a purely recurrent implementation.

These results, together with the (seemingly hidden) presence of the donut-like motif in the leading classes of decision-making models that have been used to account for categorical neural responses across brain areas and animal species [122-125, 127, 132, 137, 173-177], directly support the generality of the donut-like motif for producing categorical representations.

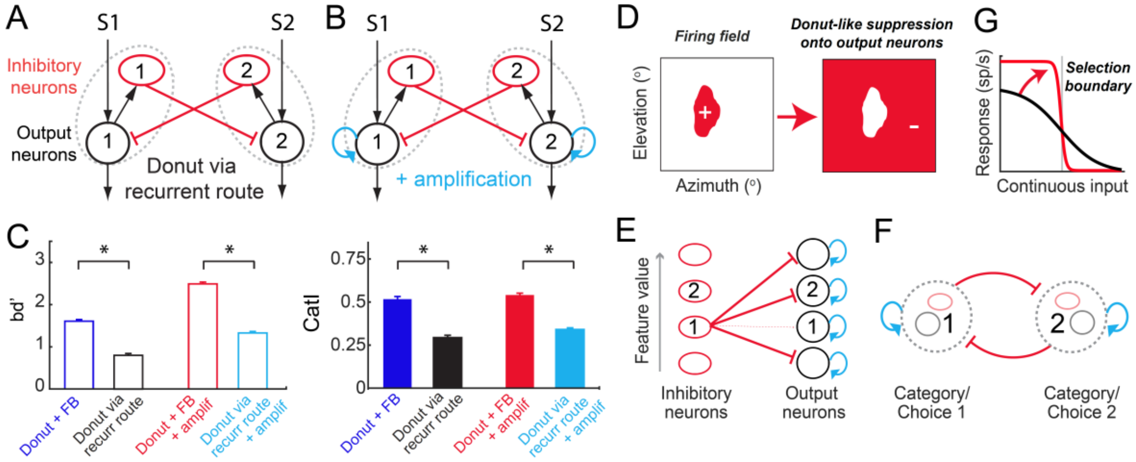


Figure 5.2: Generalizable implementation of the donut-like motif.
(A-B) Circuit with implementation of donut-like motif purely via a recurrent route, i.e., in the absence of any feedforward inhibition without (A) and with amplification (B; blue arrow); compare with, respectively, Figure 4.1CD bottom-right panels. Red ovals: inhibitory neurons, black circles: excitatory/output neurons, dashed gray ovals: populations of neurons representing each stimulus or category. **(C)** Plot of bd' (left panel; black and light blue data) and CatI (right panel; black and light blue data) computed from the responses of neuron 1 in circuits in A and B to the standard two-stimulus morphing protocol (as in Figure 4.1E). For comparison, the bd' and CatI values for donut-like motif, from Figure 4.1CD bottom-right panels, are reproduced here (dark blue and red data, respectively). '*': $p<0.05$. Left panel: black (A) vs. dark blue (Figure 4.1C-bottom–right), p = 1.7e-23; light-blue (B) vs. red (Figure 4.1D-bottom–right), p = 4.7e-34; paired t-tests with HBMC correction. Right panel: black (A) vs. dark blue (Figure 4.1C-bottom–right), p = 1.7e-25; light-blue (B) vs. red (Figure 4.1D-bottom–right), p = 2.5e-31; paired t-tests with HBMC correction. **(D-G)** Graphical summary. (D) Donut-like inhibition, i.e.,

inhibition (-) driven by preferred inputs (+) and delivered to all non-preferred inputs, can be implemented in neural circuits either in a feedforward manner (E) or in a recurrent manner (F), to generate categorical representations (G). (E, F) Red ovals: inhibitory neurons, black circles: excitatory/output neurons. E, dashed thin red arrow indicates absence of inhibitory projection. E based on Figure 4.1, Figure 4.3, Figure 4.5, Figure 4.7; F is a simplified representation of model in B, in which populations of neurons representing each category or choice (dashed grey circle) mutually inhibiting one another.

5.2 Open questions and future work

In my thesis, we have unpacked the neural mechanisms of two essential computations of stimulus selection, namely ability to select across all pairs of stimulus locations, and ability to generate categorical, robust-to-noise representations. However, selection consists of multiple computations. Some of the other computations that are essential for selection are flexibility, ability to select in the presence of multiple (>2) stimuli, unitary choice generation etc. [133]. The mechanisms of these are yet to be uncovered.

Additionally, all of our experiments were performed in head fixed barn owls passively viewing stimuli presented in front of them. Whereas, this approach has its limitations, it helped us design experiments that probe specific components of the midbrain circuit and uncover their core functions. Whether our results translate to a behaving animal in the context of a spatial attention task will need further exploration and will test the generalizability of our findings regarding circuit mechanisms to spatial attention, and to other species.

In this thesis, we directly demonstrated that the avian midbrain solves challenging computational problems using efficient algorithms. How these can help inform the design of machines and intelligent systems is a worthy question to explore.

I discuss some of these open questions and how they relate to our findings briefly in the sections below.

5.2.1 Flexibility in stimulus selection

As shown in a previously published work [94], stimulus selection in the owl OTid operates based on relative strengths of competing stimuli, and not on their absolute strengths, i.e. stimulus selection in the owl OTid is flexible. Modeling work has suggested that the most efficient circuit motif for implementing flexible stimulus selection is through feedback inhibition between inhibitory neurons encoding for competing stimuli [139]. Follow-up experiments in brain slices have identified the presence of such a motif in the avian Imc [145]. However, whether feedback inhibition between Imc neurons is indeed the functional source of flexible stimulus selection in owl midbrain selection network is yet to be determined. This would require testing the effect of inactivation of the inhibitory projections between the Imc neurons on the flexibility of stimulus selection in the OTid.

5.2.2 Multiple competing stimuli

Another important detail of our work is that it primarily focuses on competition between two competing stimuli, with one stimulus presented inside and the other outside the RF of the OTid neuron. But how would the responses of an OTid neuron be impacted by the presence of multiple competing stimuli? The expectation based on the encoding of priority in the OTid space map is that the switch value of an OTid would still depend on (and be roughly equal to) the strength of the RF stimulus (i.e. an OTid neuron would fire at a higher value when the stimulus inside its RF is the strongest among the competing stimuli, and switch to a lower spiking rate when the strength of any of the competing stimuli exceeds the strength of the RF stimulus). However, whether the mechanisms of competition translate naturally to a situation with multiple competing stimuli is still unknown and a systematic study to characterize this is under active investigation.

5.2.3 Cross-hemispheric stimulus selection

Experiments in my dissertation and previously published work have typically studied competition between stimuli located in the same hemi field. The interconnections between Imc-OT-Ipc within the same hemisphere of the brain have been fairly well characterized. However, as shown in previously published work [75], a competing stimulus presented in the opposite sensory hemifield from the RF stimulus (and therefore, encoded by the other brain hemisphere than the one containing the recording electrode), is also able to suppress the responses of an OTid neuron. Interestingly, there are no known direct projections from the Imc in one hemisphere to the OTid in the opposite hemisphere. What, then, is the source of cross-hemispheric competitive suppression? One potential candidate for cross-hemispheric transfer of information is the nucleus isthmi disseminated (ID), a glutamatergic nucleus which and receives input from the ipslateral OT and projects to the contralateral OT [79]. However, it is not an inhibitory nucleus and cannot directly send suppression to the OT in the opposite hemisphere. A potential mechanism might be for Imc to direct inhibition to OT in the other hemisphere by suppressing the output of the ipsilateral ID, thereby, reducing excitatory drive to the contralateral OT. It will be important to investigate the exact source of cross-hemispheric competitive suppression, the computational mechanisms that underlie this suppression, and to address whether the mechanisms of cross-hemispheric competition are same or different compared to those for within-hemispheric competition.

5.2.4 Mechanisms of stimulus selection within Imc

Recent studies have also suggested that Imc not only plays a key source in solving stimulus selection in the OT, but is itself a site of stimulus competition [178]. In fact, the signatures of stimulus selection appear earlier and are more precise in the Imc when compared to the OT [179]. This leads to an interesting question of whether the site of the priority map is indeed the Imc, and if the signatures of competition in the OT are a mere result of transmission of this information from the Imc. If so, then what are the mechanisms of the essential computations of stimulus selection (such as ability to select at all pairs of stimulus location, and the ability to categorize) within the Imc?

5.2.5 Information encoding in Imc

The encoding of space using multilobed Imc RFs is primarily driven by the scarcity of neurons in the coronal planes of Imc (section 2.3). As we suggested in Chapter 3, the scarcity of the neurons in the Imc could be a consequence of minimizing a combination of circuit wiring and metabolic costs (Figure 3.6). However, how multilobed RFs impact the capacity and accuracy of information encoding in the midbrain network in comparison to single lobed RFs is an important question that remains to be answered. Tuning width and accuracy (as measured using a fisher information metric) have been shown to depend critically on the dimension of the space being encoded [180]. It has been shown that for one-dimensional encoding, neurons with shaper tuning curves encode more information in comparison to those with broader tuning curves. This is consistent with our discovery of multliobed RFs along the elevational space in contrast to previous reports of elongated RFs [105]. However, detailed analyses that take into account the physiological properties of the Imc, and its functional connectivity with the OT and Ipc are essential to unpack the relation between Imc RFs and the information being encoded.

5.2.6 Mechanisms of stimulus selection for spatial attention (i.e., in behaving animals)

The midbrain attention network is a conserved structure across species. The well-organized midbrain attention network in barn owls together with their multisensory expertise are major advantages of using owl as a model to study mechanisms of stimulus selection. However, using them for trained behavior such as in a task for spatial attention, while not impossible, is highly challenging. Secondly, as they are not mammals, the question of whether the mechanisms uncovered here generalize to mammalian species is currently open. Thus, the use of a mammalian species that can be trained on spatial attention tasks would allow the testing of whether the findings from this work generalize to the control of spatial attention in mammals. In mammals, the less well-organized midbrain necessitates the use of modern genetics-based tools such as calcium imaging, optogenetics, and chemogenetics for the study of neural function underlying selection in a cell-type specific manner. For the use of such tools, including the use of transgenic lines, mice stand out as a powerful candidate mammalian model.

Investigations of whether multilobed RFs also exist in the mouse Imc (pLTN), and does the mouse pLTN also use a combinatorial neural inhibition mechanism to solve location-invariant stimulus selection are potentially very exciting. Additionally, if the pLTN does have multilobed RFs, does it also follow the organizational structure that is similar to the owl Imc, or are there species-specific differences that relate to the differences in the behavioral mechanisms essential for survival in these two animals (an owl is a predator that primarily spends its time at higher elevations whereas a mouse is a prey that spends its time on the ground).

Similarly, does the donut-like circuit motif also exist in the mouse midbrain, and if it does, is it also essential for generating categorical responses like in the owl OT? If the donut-like circuit motif exists in the mouse midbrain, will manipulating the donut (like we did in the owl setup; Figure 4.7) result in predictable deficits in the mouse's ability to select the stronger of two competing stimuli in the presence of sensory noise?

A final set of critical questions that can then be addressed is how the different components of the midbrain selection network contribute to attentional behavior. As demonstrated recently, mice are able to pay attention and select relevant targets in the presence of distractors [159, 181]. Future studies could directly test whether manipulating activity in the Imc results in predictable changes in behavior that are consistent with the role of Imc in distractor suppression and target selection. At the same time, these techniques could also help uncover, mechanistically, how specific components of this midbrain network connect to forebrain areas, and how they together orchestrate the computations that help select relevant stimuli for paying attention in a dynamically changing world.

5.2.7 Relation to deep learning and AI systems

In the past decade, the fields of machine learning, AI and deep learning have exploded. The state-of-the-art performances of the deep learning systems on various recognition and classification tasks today far exceed those of algorithms from 10 years ago (for instance the recognition accuracy on the ImageNet dataset has improved from 51% to 88.4% in the last 10 years). Deep learning systems have been able to achieve human level performance in various tasks and have also been able to beat humans in fairly complicated games such as alphago [182].

Despite such tremendous progress, deep learning systems are still far from achieving human level AI. Deep networks often require large labeled data to be trained and are susceptible to being tricked, and often make ‘silly’ errors [183, 184], something that human brains are very good at avoiding. Incorporating fundamental principles of biological brain function such as attention can potentially help neural networks overcome some of these challenges. For instance, teaching a neural network as to which piece of information it needs to focus on for learning a task results in performance improvements, and may be possible to achieve with a smaller training dataset.

Given how closely linked machine learning and neuroscience are, any understanding of elemental rules underlying how the brain implements computations can inform the design of deep learning systems. It is no surprise that the design of machines can benefit from an understanding of how the brain implements computations. For instance, in recent work, we showed that key filtering techniques that are dominant among a population of neurons in the primary auditory cortex of ferrets are consistent with previously used techniques that have been shown to result in improved recognition of speech [185, 186].

In this context, questions that merit further exploration are: (a) What key aspects of the findings reported in this thesis (Chapters 2, 3 and 4) need to be incorporated into the design of an artificial system that relies on the selection of the most important piece of information among multiple competing options (for example search and navigation, driverless cars, etc) for improved function? (b) How might the combinatorial inhibition mechanism and the donut-like motif help implement selection in a deep network to allow it to 'pay attention' effectively?

5.2.8 Attentional disorders

One of the symptoms of conditions such as ADHD, autism and schizophrenia, is a dysfunction in the ability to select and pay attention to relevant information. Current diagnosis typically relies on observation and self-report of behavior [187]. Current treatments involve medications (for instance, stimulants) that do not target specific brain regions, but rather act on large parts of the brain, primarily because a clear understanding of various the components of these complex disorders is lacking. Dissecting the neural circuit mechanisms of selective attention can directly inform hypotheses regarding which neural mechanisms of selective attention (the computations, and also the anatomical and functional sources) affect what aspects of attentional dysfunction in these disorders. This can lead to better diagnosis as well as targeted treatments of these neural disorders that affect millions of families around the world.

6. Bibliography

[1] K. Koch *et al.*, "How much the eye tells the brain," (in English), *Current Biology,* vol. 16, no. 14, pp. 1428-1434, Jul 25 2006.

[2] D. Attwell and S. B. Laughlin, "An energy budget for signaling in the grey matter of the brain," (in English), *Journal of Cerebral Blood Flow and Metabolism,* vol. 21, no. 10, pp. 1133-1145, Oct 2001.

[3] B. L. Chen, D. H. Hall, and D. B. Chklovskii, "Wiring optimization can relate neuronal structure and function," *Proc Natl Acad Sci U S A,* vol. 103, no. 12, pp. 4723-8, Mar 21 2006.

[4] H. B. Barlow, "Single units and sensation: a neuron doctrine for perceptual psychology?," *Perception,* vol. 1, no. 4, pp. 371-94, 1972.

[5] P. Lennie, "The cost of cortical computation," *Curr Biol,* vol. 13, no. 6, pp. 493-7, Mar 18 2003.

[6] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *Psychological Review,* vol. 63, no. 2, pp. 81-97, 1956.

[7] D. E. Broadbent, "Perception and Communication," *Oxford University Press.,* 1958.

[8] N. P. Moray, "Attention in dichotic listening: Affective cues and the influence of instructions," *Quarterly Journal of Experimental Psychology,* vol. 11, pp. 56–60, 1959.

[9] C. E. Cherry, "Some experiments on the recognition of speech with one and with two ears," *J. acoust. Soc. Amer.,* vol. 25, pp. 975-979, 1953.

[10] A. M. Treisman, "Contextual cues in selective listening," *Quart. Y. exp. Psychol.,* vol. 12, pp. 242-248, 1960.

[11] A. M. Treisman, "Selective Attention in Man," (in eng), *Br Med Bull,* vol. 20, pp. 12-6, Jan 1964.

[12] M. Carrasco, "Visual attention: the past 25 years," (in eng), *Vision Res,* vol. 51, no. 13, pp. 1484-525, Jul 1 2011.

[13] E. Kowler, "Eye movements: the past 25 years," *Vision Res,* vol. 51, no. 13, pp. 1457-83, Jul 1 2011.

[14] K. Nakayama and P. Martini, "Situating visual search," *Vision Res,* vol. 51, no. 13, pp. 1526-37, Jul 1 2011.

[15] K. A. MacLean, S. R. Aichele, D. A. Bridwell, G. P. Mangun, E. Wojciulik, and C. D. Saron, "Interactions between endogenous and exogenous attention during vigilance," (in English), *Attention Perception & Psychophysics,* vol. 71, no. 5, pp. 1042-1058, Jul 2009.

[16] C. L. Folk, A. B. Leber, and H. E. Egeth, "Made you blink! Contingent attentional capture produces a spatial blink," (in English), *Perception & Psychophysics,* vol. 64, no. 5, pp. 741-753, Jul 2002.

[17] C. J. Leonard, A. Balestreri, and S. J. Luck, "Interactions Between Space-Based and Feature-Based Attention," (in English), *Journal of Experimental Psychology-Human Perception and Performance,* vol. 41, no. 1, pp. 11-16, Feb 2015.

[18] C. M. A. Darwin, "On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life " *London: John Murray,* p. p. 502, 1859, retrieved 1 March 2011.

[19] J. H. Reynolds and L. Chelazzi, "Attentional modulation of visual processing," (in eng), *Annu Rev Neurosci,* vol. 27, pp. 611-47, 2004.

[20] M. E. Goldberg and R. H. Wurtz, "Activity of superior colliculus in behaving monkey. II. Effect of attention on neuronal responses," (in eng), *J Neurophysiol,* vol. 35, no. 4, pp. 560-74, Jul 1972.

[21] S. Treue and J. H. Maunsell, "Attentional modulation of visual motion processing in cortical areas MT and MST," *Nature,* vol. 382, no. 6591, pp. 539-41, Aug 8 1996.

[22] J. H. Reynolds, T. Pasternak, and R. Desimone, "Attention increases sensitivity of V4 neurons," *Neuron,* vol. 26, no. 3, pp. 703-14, Jun 2000.

[23] J. W. Bisley and M. E. Goldberg, "Neuronal activity in the lateral intraparietal area and spatial attention," *Science,* vol. 299, no. 5603, pp. 81-6, Jan 3 2003.

[24] R. M. McPeek and E. L. Keller, "Saccade target selection in the superior colliculus during a visual search task," (in eng), *J Neurophysiol,* vol. 88, no. 4, pp. 2019-34, Oct 2002.

[25] A. Thiele *et al.*, "Attention Induced Gain Stabilization in Broad and Narrow-Spiking Cells in the Frontal Eye-Field of Macaque Monkeys," (in English), *Journal of Neuroscience,* vol. 36, no. 29, pp. 7601-7612, Jul 20 2016.

[26] C. J. McAdams and J. H. R. Maunsell, "Effects of attention on the reliability of individual neurons in monkey visual cortex," (in English), *Neuron,* vol. 23, no. 4, pp. 765-773, Aug 1999.

[27] J. F. Mitchell, K. A. Sundberg, and J. H. Reynolds, "Differential Attention-Dependent Response Modulation across Cell Classes in Macaque Visual Area V4," *Neuron,* vol. 55, no. 1, pp. 131-41, Jul 5 2007.

[28] M. R. Cohen and J. H. Maunsell, "Attention improves performance primarily by reducing interneuronal correlations," *Nat Neurosci,* vol. 12, no. 12, pp. 1594-600, Dec 2009.

[29] J. F. Mitchell, K. A. Sundberg, and J. H. Reynolds, "Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4," *Neuron,* vol. 63, no. 6, pp. 879-88, Sep 24 2009.

[30] L. F. Abbott and P. Dayan, "The effect of correlated variability on the accuracy of a population code," *Neural Comput,* vol. 11, no. 1, pp. 91-101, Jan 1 1999.

[31] B. B. Averbeck, P. E. Latham, and A. Pouget, "Neural correlations, population coding and computation," *Nat Rev Neurosci,* vol. 7, no. 5, pp. 358-66, May 2006.

[32] P. N. Steinmetz, A. Roy, P. J. Fitzgerald, S. S. Hsiao, K. O. Johnson, and E. Niebur, "Attention modulates synchronized neuronal firing in primate somatosensory cortex," *Nature,* vol. 404, no. 6774, pp. 187-90, Mar 9 2000.

[33] P. Fries, J. H. Reynolds, A. E. Rorie, and R. Desimone, "Modulation of oscillatory neuronal synchronization by selective visual attention," (in eng), *Science,* vol. 291, no. 5508, pp. 1560-3, Feb 23 2001.

[34] E. Salinas and T. J. Sejnowski, "Impact of correlated synaptic input on output firing rate and variability in simple neuronal models," *J Neurosci,* vol. 20, no. 16, pp. 6193-209, Aug 15 2000.

[35] A. Bollimunta, J. Mo, C. E. Schroeder, and M. Ding, "Neuronal mechanisms and attentional modulation of corticothalamic alpha oscillations," *J Neurosci,* vol. 31, no. 13, pp. 4935-43, Mar 30 2011.

[36] G. G. Gregoriou, S. J. Gotts, H. Zhou, and R. Desimone, "High-frequency, long-range coupling between prefrontal and visual cortex during attention," (in eng), *Science,* vol. 324, no. 5931, pp. 1207-10, May 29 2009.

[37] W. Klimesch, "alpha-band oscillations, attention, and controlled access to stored information," *Trends Cogn Sci,* vol. 16, no. 12, pp. 606-17, Dec 2012.

[38] B. F. Handel, T. Haarmeier, and O. Jensen, "Alpha oscillations correlate with the successful inhibition of unattended stimuli," *J Cogn Neurosci,* vol. 23, no. 9, pp. 2494-502, Sep 2011.

[39] G. G. Gregoriou, S. Paneri, and P. Sapountzis, "Oscillatory synchrony as a mechanism of attentional processing," (in English), *Brain Research,* vol. 1626, pp. 165-182, Nov 11 2015.

[40] A. K. Engel, A. K. Kreiter, P. Konig, and W. Singer, "Synchronization of Oscillatory Neuronal Responses between Striate and Extrastriate Visual Cortical Areas of the Cat," (in English), *Proceedings of the National Academy of Sciences of the United States of America,* vol. 88, no. 14, pp. 6048-6052, Jul 1991.

[41] P. R. Roelfsema, A. K. Engel, P. Konig, and W. Singer, "Visuomotor integration is associated with zero time-lag synchronization among cortical areas," (in English), *Nature,* vol. 385, no. 6612, pp. 157-161, Jan 9 1997.

[42] J. Csicsvari, B. Jamieson, K. D. Wise, and G. Buzsaki, "Mechanisms of gamma oscillations in the hippocampus of the behaving rat," *Neuron,* vol. 37, no. 2, pp. 311-22, Jan 23 2003.

[43] T. J. Buschman and E. K. Miller, "Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices," *Science,* vol. 315, no. 5820, pp. 1860-2, Mar 30 2007.

[44] G. G. Gregoriou, A. F. Rossi, L. G. Ungerleider, and R. Desimone, "Lesions of prefrontal cortex reduce attentional modulation of neuronal responses and synchrony in V4," *Nat Neurosci,* vol. 17, no. 7, pp. 1003-11, Jul 2014.

[45] D. Sridharan, K. Boahen, and E. I. Knudsen, "Space coding by gamma oscillations in the barn owl optic tectum," *J Neurophysiol,* vol. 105, no. 5, pp. 2005-17, May 2011.

[46] S. Ray and J. H. Maunsell, "Different origins of gamma rhythm and high-gamma activity in macaque visual cortex," *PLoS Biol,* vol. 9, no. 4, p. e1000610, Apr 2011.

[47] M. F. Carr, M. P. Karlsson, and L. M. Frank, "Transient slow gamma synchrony underlies hippocampal memory replay," *Neuron,* vol. 75, no. 4, pp. 700-13, Aug 23 2012.

[48] T. Gruber, A. Keil, and M. M. Muller, "Modulation of induced gamma band responses and phase synchrony in a paired associate learning task in the human EEG," *Neurosci Lett,* vol. 316, no. 1, pp. 29-32, Dec 4 2001.

[49] C. Koch and S. Ullman, "Shifts in selective visual attention: towards the underlying neural circuitry," *Hum Neurobiol,* vol. 4, no. 4, pp. 219-27, 1985.

[50] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 20, no. 11, pp. 1254-1259, 1998.

[51] M. I. Posner and Y. Cohen, *Components of performance. In: Attention and performance*. Hillsdale: Erlbaum, 1984, pp. 531-556.

[52] H. Kwak and H. Egeth, "Consequences of allocating attention to locations and to other attributes," *Perception & Psychophysics,* vol. 51, pp. 455-464, 19992.

[53] R. M. Klein, "Inhibition of return," *Trends Cogn Sci,* vol. 4, no. 4, pp. 138-147, Apr 2000.

[54] C. Kayser, C. I. Petkov, M. Lippert, and N. K. Logothetis, "Mechanisms for allocating auditory attention: an auditory saliency map," *Curr Biol,* vol. 15, no. 21, pp. 1943-7, Nov 8 2005.

[55] E. M. Kaya and M. Elhilali, "Modelling auditory attention," *Philos Trans R Soc Lond B Biol Sci,* vol. 372, no. 1714, Feb 19 2017.

[56] E. M. Kaya and M. Elhilali, "Investigating bottom-up auditory attention," *Front Hum Neurosci,* vol. 8, p. 327, 2014.

[57] L. Itti and C. Koch, "Computational modelling of visual attention," *Nat Rev Neurosci,* vol. 2, no. 3, pp. 194-203, Mar 2001.

[58] E. Niebur and C. Koch, "Control of Selective Visual Attention: Modeling the "Where" Pathway " *Advances in Neural Information Processing Systems (NIPS),* pp. 802-808, 1996.

[59] J. H. Fecteau and D. P. Munoz, "Salience, relevance, and firing: a priority map for target selection," (in eng), *Trends Cogn Sci,* vol. 10, no. 8, pp. 382-90, Aug 2006.

[60] Y. L. Yu, G. K. I. Mann, and R. G. Gosine, "Target Tracking for Moving Robots Using Object-based Visual Attention," (in English), *Ieee/Rsj 2010 International Conference on Intelligent Robots and Systems (Iros 2010),* 2010.

[61] T. V. Nguyen, Z. Song, and S. C. Yan, "STAP: Spatial-Temporal Attention-Aware Pooling for Action Recognition," (in English), *Ieee Transactions on Circuits and Systems for Video Technology,* vol. 25, no. 1, pp. 77-86, Jan 2015.

[62] D. Parkhurst, K. Law, and E. Niebur, "Modeling the role of salience in the allocation of overt visual attention," (in eng), *Vision Res,* vol. 42, no. 1, pp. 107-23, Jan 2002.

[63] M. Mishkin, L. G. Ungerleider, and K. A. Macko, "Object Vision and Spatial Vision - 2 Cortical Pathways," (in English), *Trends in Neurosciences,* vol. 6, no. 10, pp. 414-417, 1983.

[64] R. Veale, Z. M. Hafed, and M. Yoshida, "How is visual salience computed in the brain? Insights from behaviour, neurobiology and modelling," *Philos Trans R Soc Lond B Biol Sci,* vol. 372, no. 1714, Feb 19 2017.

[65] R. H. Wurtz and J. E. Albano, "Visual-motor function of the primate superior colliculus," (in eng), *Annu Rev Neurosci,* vol. 3, pp. 189-226, 1980.

[66] A. S. Mitchell, S. M. Sherman, M. A. Sommer, R. G. Mair, R. P. Vertes, and Y. Chudasama, "Advances in understanding mechanisms of thalamic relays in cognition and behavior," *J Neurosci,* vol. 34, no. 46, pp. 15340-6, Nov 12 2014.

[67] S. M. Sherman, "Thalamus plays a central role in ongoing cortical functioning," *Nat Neurosci,* vol. 19, no. 4, pp. 533-41, Apr 2016.

[68] F. Crick, "Function of the thalamic reticular complex: the searchlight hypothesis," *Proc Natl Acad Sci U S A,* vol. 81, no. 14, pp. 4586-90, Jul 1984.

[69] K. McAlonan, J. Cavanaugh, and R. H. Wurtz, "Guarding the gateway to cortex with attention in visual thalamus," *Nature,* vol. 456, no. 7220, pp. 391-4, Nov 20 2008.

[70] D. H. O'Connor, M. M. Fukui, M. A. Pinsk, and S. Kastner, "Attention modulates responses in the human lateral geniculate nucleus," *Nat Neurosci,* vol. 5, no. 11, pp. 1203-9, Nov 2002.

[71] R. D. Wimmer, L. I. Schmitt, T. J. Davidson, M. Nakajima, K. Deisseroth, and M. M. Halassa, "Thalamic control of sensory selection in divided attention," *Nature,* vol. 526, no. 7575, pp. 705-9, Oct 29 2015.

[72] K. R. Gegenfurtner and D. C. Kiper, "Color vision," *Annu Rev Neurosci,* vol. 26, pp. 181-206, 2003.

[73] K. G. Thompson and N. P. Bichot, "A visual salience map in the primate frontal eye field," (in eng), *Prog Brain Res,* vol. 147, pp. 251-62, 2005.

[74] J. P. Gottlieb, M. Kusunoki, and M. E. Goldberg, "The representation of visual salience in monkey parietal cortex," (in eng), *Nature,* vol. 391, no. 6666, pp. 481-4, Jan 29 1998.

[75] S. P. Mysore, A. Asadollahi, and E. I. Knudsen, "Global inhibition and stimulus competition in the owl optic tectum," (in eng), *J Neurosci,* vol. 30, no. 5, pp. 1727-38, Feb 3 2010.

[76] M. E. Goldberg and R. H. Wurtz, "Activity of superior colliculus in behaving monkey. I. Visual receptive fields of single neurons," (in eng), *J Neurophysiol,* vol. 35, no. 4, pp. 542-59, Jul 1972.

[77] E. I. Knudsen, "Auditory and visual maps of space in the optic tectum of the owl," *J. Neurosci.,* vol. 2, pp. 1177-1194, 1982.

[78] J. K. Harting, "Descending pathways from the superior collicullus: an autoradiographic analysis in the rhesus monkey (Macaca mulatta)," (in eng), *J Comp Neurol,* vol. 173, no. 3, pp. 583-612, Jun 1 1977.

[79] E. I. Knudsen, "Control from below: the role of a midbrain network in spatial attention," (in eng), *Eur J Neurosci,* vol. 33, no. 11, pp. 1961-72, Jun 2011.

[80] U. C. Drager and D. H. Hubel, "Responses to Visual Stimulation and Relationship between Visual, Auditory, and Somatosensory Inputs in Mouse Superior Colliculus," (in English), *Journal of Neurophysiology,* vol. 38, no. 3, pp. 690-713, 1975.

[81] S. E. Boehnke and D. P. Munoz, "On the importance of the transient visual response in the superior colliculus," (in eng), *Curr Opin Neurobiol,* vol. 18, no. 6, pp. 544-51, Dec 2008.

[82] H. J. Karten, W. Hodos, W. J. Nauta, and A. M. Revzin, "Neural connections of the "visual wulst" of the avian telencephalon. Experimental studies in the piegon (Columba livia) and owl (Speotyto cunicularia)," *J Comp Neurol,* vol. 150, no. 3, pp. 253-78, Aug 1973.

[83] E. I. Knudsen, Y. E. Cohen, and T. Masino, "Characterization of a forebrain gaze field in the archistriatum of the barn owl: microstimulation and anatomical connections," *J Neurosci,* vol. 15, no. 7 Pt 2, pp. 5139-51, Jul 1995.

[84] E. I. Knudsen and P. F. Knudsen, "Space-Mapped Auditory Projections from the Inferior Colliculus to the Optic Tectum in the Barn Owl (Tyto-Alba)," (in English), *Journal of Comparative Neurology,* vol. 218, no. 2, pp. 187-196, 1983.

[85] Y. Wang, D. E. Major, and H. J. Karten, "Morphology and connections of nucleus isthmi pars magnocellularis in chicks (Gallus gallus)," *J Comp Neurol,* vol. 469, no. 2, pp. 275-97, Feb 2 2004.

[86] Y. Wang, H. Luksch, N. C. Brecha, and H. J. Karten, "Columnar projections from the cholinergic nucleus isthmi to the optic tectum in chicks (Gallus gallus): a possible substrate for synchronizing tectal channels," *J Comp Neurol,* vol. 494, no. 1, pp. 7-35, Jan 1 2006.

[87] G. Marin, C. Salas, E. Sentis, X. Rojas, J. C. Letelier, and J. Mpodozis, "A cholinergic gating mechanism controlled by competitive interactions in the optic tectum of the pigeon," *J Neurosci,* vol. 27, no. 30, pp. 8112-21, Jul 25 2007.

[88] G. Marin, J. Mpodozis, E. Sentis, T. Ossandon, and J. C. Letelier, "Oscillatory bursts in the optic tectum of birds represent re-entrant signals from the nucleus isthmi pars parvocellularis," (in eng), *J Neurosci,* vol. 25, no. 30, pp. 7081-9, Jul 27 2005.

[89] C. D. Carello and R. J. Krauzlis, "Manipulating intent: evidence for a causal role of the superior colliculus in target selection," *Neuron,* vol. 43, no. 4, pp. 575-83, Aug 19 2004.

[90] J. Cavanaugh and R. H. Wurtz, "Subcortical modulation of attention counters change blindness," *J Neurosci,* vol. 24, no. 50, pp. 11236-43, Dec 15 2004.

[91] J. R. Muller, M. G. Philiastides, and W. T. Newsome, "Microstimulation of the superior colliculus focuses attention without moving the eyes," *Proc Natl Acad Sci U S A,* vol. 102, no. 3, pp. 524-9, Jan 18 2005.

[92] R. M. McPeek and E. L. Keller, "Deficits in saccade target selection after inactivation of superior colliculus," *Nat Neurosci,* vol. 7, no. 7, pp. 757-63, Jul 2004.

[93] L. P. Lovejoy and R. J. Krauzlis, "Inactivation of primate superior colliculus impairs covert selection of signals for perceptual judgments," (in eng), *Nat Neurosci,* vol. 13, no. 2, pp. 261-6, Feb 2010.

[94] S. P. Mysore, A. Asadollahi, and E. I. Knudsen, "Signaling of the strongest stimulus in the owl optic tectum," (in eng), *J Neurosci,* vol. 31, no. 14, pp. 5186-96, Apr 6 2011.

[95] S. P. Mysore and E. I. Knudsen, "Flexible categorization of relative stimulus strength by the optic tectum," (in eng), *J Neurosci,* vol. 31, no. 21, pp. 7745-52, May 25 2011.

[96] P. Lee and W. C. Hall, "An in vitro study of horizontal connections in the intermediate layer of the superior colliculus," *J Neurosci,* vol. 26, no. 18, pp. 4763-8, May 3 2006.

[97] T. Isa and W. C. Hall, "Exploring the superior colliculus in vitro," (in eng), *J Neurophysiol,* vol. 102, no. 5, pp. 2581-93, Nov 2009.

[98] A. M. Graybiel, "A satellite system of the superior colliculus: the parabigeminal nucleus and its projections to the superficial collicular layers," (in eng), *Brain Res,* vol. 145, no. 2, pp. 365-74, Apr 28 1978.

[99] M. I. Sereno and P. S. Ulinski, "Caudal topographic nucleus isthmi and the rostral nontopographic nucleus isthmi in the turtle, Pseudemys scripta," (in eng), *J Comp Neurol,* vol. 261, no. 3, pp. 319-46, Jul 15 1987.

[100] S. P. Mysore and E. I. Knudsen, "A shared inhibitory circuit for both exogenous and endogenous control of stimulus selection," *Nat Neurosci,* vol. 16, no. 4, pp. 473-8, Apr 2013.

[101] D. E. Winkowski and E. I. Knudsen, "Top-down gain control of the auditory space map by gaze control circuitry in the barn owl," *Nature,* vol. 439, no. 7074, pp. 336-9, Jan 19 2006.

[102] T. Moore and K. M. Armstrong, "Selective gating of visual signals by microstimulation of frontal cortex," *Nature,* vol. 421, no. 6921, pp. 370-3, Jan 23 2003.

[103] K. A. Maczko, P. F. Knudsen, and E. I. Knudsen, "Auditory and visual space maps in the cholinergic nucleus isthmi pars parvocellularis of the barn owl," *J Neurosci,* vol. 26, no. 49, pp. 12799-806, Dec 6 2006.

[104] A. Asadollahi, S. P. Mysore, and E. I. Knudsen, "Stimulus-driven competition in a cholinergic midbrain nucleus," (in eng), *Nat Neurosci,* vol. 13, no. 7, pp. 889-95, Jul 2010.

[105] D. P. Li, Q. Xiao, and S. R. Wang, "Feedforward construction of the receptive field and orientation selectivity of visual neurons in the pigeon," *Cereb Cortex,* vol. 17, no. 4, pp. 885-93, Apr 2007.

[106] Y. C. Wang and B. J. Frost, "Visual response characteristics of neurons in the nucleus isthmi magnocellularis and nucleus isthmi parvocellularis of pigeons," *Exp Brain Res,* vol. 87, no. 3, pp. 624-33, 1991.

[107] G. Rizzolatti, R. Camarda, L. A. Grupp, and M. Pisa, "Inhibitory effect of remote visual stimuli on visual responses of cat superior colliculus: spatial and temporal factors," (in eng), *J Neurophysiol,* vol. 37, no. 6, pp. 1262-75, Nov 1974.

[108] C. A. Goddard, J. Huguenard, and E. Knudsen, "Parallel midbrain microcircuits perform independent temporal transformations," *J Neurosci,* vol. 34, no. 24, pp. 8130-8, Jun 11 2014.

[109] A. Rodriguez and A. Laio, "Machine learning. Clustering by fast search and find of density peaks," *Science,* vol. 344, no. 6191, pp. 1492-6, Jun 27 2014.

[110] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the number of clusters in a dataset via the gap statistic," *J. R. Statist. Soc. B,* vol. 63, no. 2, pp. 411-423, 2001.

[111] M. A. Garcia-Cabezas, Y. J. John, H. Barbas, and B. Zikopoulos, "Distinction of Neurons, Glia and Endothelial Cells in the Cerebral Cortex: An Algorithm Based on Cytological Features," *Front Neuroanat,* vol. 10, p. 107, 2016.

[112] M. S. Fee, P. P. Mitra, and D. Kleinfeld, "Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability," (in eng), *J Neurosci Methods,* vol. 69, no. 2, pp. 175-88, Nov 1996.

[113] A. D. S. Bala, M. W. Spitzer, and T. T. Takahashi, "Prediction of auditory spatial acuity from neural images on the owl's auditory space map," *Nature,* vol. 424, p. 4, 14th August 2003 2003.

[114] M. Carrasco and K. S. Frieder, "Cortical magnification neutralizes the eccentricity effect in visual search," *Vision Res,* vol. 37, no. 1, pp. 63-82, Jan 1997.

[115] M. Schlueter, M. Gerdts, and J. J. Ruckmann, "A numerical study of MIDACO on 100 MINLP benchmarks," *Optim,* vol. 61, no. 7, pp. 873-900, 2012.

[116] C. Lee, W. H. Rohrer, and D. L. Sparks, "Population coding of saccadic eye movements by neurons in the superior colliculus," *Nature,* vol. 332, no. 6162, pp. 357-60, Mar 24 1988.

[117] J. E. Lewis and W. B. Kristan, Jr., "A neuronal network for computing population vectors in the leech," *Nature,* vol. 391, no. 6662, pp. 76-9, Jan 01 1998.

[118] A. P. Georgopoulos, A. B. Schwartz, and R. E. Kettner, "Neuronal population coding of movement direction," *Science,* vol. 233, no. 4771, pp. 1416-9, Sep 26 1986.

[119] W. J. Ma, J. M. Beck, P. E. Latham, and A. Pouget, "Bayesian inference with probabilistic population codes," (in eng), *Nat Neurosci,* vol. 9, no. 11, pp. 1432-8, Nov 2006.

[120] G. Sicard and A. Holley, "Receptor cell responses to odorants: similarities and differences among odorants," *Brain Res,* vol. 292, no. 2, pp. 283-96, Feb 06 1984.

[121] M. Yilmaz and M. Meister, "Rapid innate defensive responses of mice to looming visual stimuli," *Curr Biol,* vol. 23, no. 20, pp. 2011-5, Oct 21 2013.

[122] X. Jiang, M. A. Chevillet, J. P. Rauschecker, and M. Riesenhuber, "Training Humans to Categorize Monkey Calls: Auditory Feature- and Category-Selective Neural Tuning Changes," *Neuron,* vol. 98, no. 2, pp. 405-416 e4, Apr 18 2018.

[123] D. J. Freedman and J. A. Assad, "Experience-dependent representation of visual categories in parietal cortex," (in eng), *Nature,* vol. 443, no. 7107, pp. 85-8, Sep 7 2006.

[124] D. J. Freedman, M. Riesenhuber, T. Poggio, and E. K. Miller, "Categorical representation of visual stimuli in the primate prefrontal cortex," (in eng), *Science,* vol. 291, no. 5502, pp. 312-6, Jan 12 2001.

[125] J. Niessing and R. W. Friedrich, "Olfactory pattern classification by discrete neuronal network states," (in eng), *Nature,* vol. 465, no. 7294, pp. 47-52, May 6 2010.

[126] J. Hirokawa, A. Vaughan, P. Masset, T. Ott, and A. Kepecs, "Frontal cortex neuron types categorically encode single decision variables," *Nature,* vol. 576, no. 7787, pp. 446-451, Dec 2019.

[127] B. Bathellier, L. Ushakova, and S. Rumpel, "Discrete neocortical dynamics predict behavioral categorization of sounds," *Neuron,* vol. 76, no. 2, pp. 435-49, Oct 18 2012.

[128] Y. Xin, L. Zhong, Y. Zhang, T. Zhou, J. Pan, and N. L. Xu, "Sensory-to-Category Transformation via Dynamic Reorganization of Ensemble Structures in Mouse Auditory Cortex," *Neuron,* vol. 103, no. 5, pp. 909-921 e6, Sep 4 2019.

[129] T. Jovanic *et al.*, "Competitive Disinhibition Mediates Behavioral Choice and Sequences in Drosophila," *Cell,* vol. 167, no. 3, pp. 858-870 e19, Oct 20 2016.

[130] T. A. Engel, W. Chaisangmongkon, D. J. Freedman, and X. J. Wang, "Choice-correlated activity fluctuations underlie learning of neuronal category representation," *Nat Commun,* vol. 6, p. 6454, Mar 11 2015.

[131] X. J. Wang, "Neural dynamics and circuit mechanisms of decision-making," *Curr Opin Neurobiol,* vol. 22, no. 6, pp. 1039-46, Dec 2012.

[132] C. K. Machens, R. Romo, and C. D. Brody, "Flexible control of mutual inhibition: a neural model of two-interval discrimination," (in eng), *Science,* vol. 307, no. 5712, pp. 1121-4, Feb 18 2005.

[133] S. P. Mysore and N. B. Kothari, "Mechanisms of competitive selection: A canonical neural circuit framework " *In review*

[134] S. R. Olsen, V. Bhandawat, and R. I. Wilson, "Divisive normalization in olfactory population codes," (in eng), *Neuron,* vol. 66, no. 2, pp. 287-99, Apr 29 2010.

[135] A. Bollimunta and J. Ditterich, "Local computation of decision-relevant net sensory evidence in parietal cortex," *Cereb Cortex,* vol. 22, no. 4, pp. 903-17, Apr 2012.

[136] H. K. Hartline, H. G. Wagner, and F. Ratliff, "Inhibition in the eye of Limulus," (in eng), *J Gen Physiol,* vol. 39, no. 5, pp. 651-73, May 20 1956.

[137] X. J. Wang, "Decision making in recurrent neuronal circuits," (in eng), *Neuron,* vol. 60, no. 2, pp. 215-34, Oct 23 2008.

[138] Q. Chen, Z. Pei, D. Koren, and W. Wei, "Stimulus-dependent recruitment of lateral inhibition underlies retinal direction selectivity," *Elife,* vol. 5, Dec 8 2016.

[139] S. P. Mysore and E. I. Knudsen, "Reciprocal inhibition of inhibition: a circuit motif for flexible categorization in stimulus selection," (in eng), *Neuron,* vol. 73, no. 1, pp. 193-205, Jan 12 2012.

[140] X. J. Wang, "Probabilistic decision making by slow reverberation in cortical circuits," (in eng), *Neuron,* vol. 36, no. 5, pp. 955-68, Dec 5 2002.

[141] L. Albantakis and G. Deco, "The encoding of alternatives in multiple-choice decision making," *Proc Natl Acad Sci U S A,* vol. 106, no. 25, pp. 10308-13, Jun 23 2009.

[142] A. Asadollahi and E. I. Knudsen, "Spatially precise visual gain control mediated by a cholinergic circuit in the midbrain attention network," *Nat Commun,* vol. 7, p. 13472, Nov 17 2016.

[143] S. P. Mysore and E. I. Knudsen, "The role of a midbrain network in competitive stimulus selection," (in Eng), *Curr Opin Neurobiol,* vol. 21, no. 4, pp. 653-60, Jun 20 2011.

[144] S. P. Mysore and E. I. Knudsen, "Descending control of neural bias and selectivity in a spatial attention network: rules and mechanisms," *Neuron,* vol. 84, no. 1, pp. 214-26, Oct 1 2014.

[145] C. A. Goddard, S. P. Mysore, A. S. Bryant, J. R. Huguenard, and E. I. Knudsen, "Spatially Reciprocal Inhibition of Inhibition within a Stimulus Selection Network in the Avian Midbrain," *PLoS One,* vol. 9, no. 1, p. e85865, 2014.

[146] N. R. Mahajan and S. P. Mysore, "Combinatorial Neural Inhibition for Stimulus Selection across Space," *Cell Rep,* vol. 25, no. 5, pp. 1158-1170 e9, Oct 30 2018.

[147] M. Carandini, D. J. Heeger, and J. A. Movshon, "Linearity and normalization in simple cells of the macaque primary visual cortex," (in eng), *J Neurosci,* vol. 17, no. 21, pp. 8621-44, Nov 1 1997.

[148] M. Kouh and T. Poggio, "A canonical neural circuit for cortical nonlinear operations," *Neural Comput,* vol. 20, no. 6, pp. 1427-51, Jun 2008.

[149] K. Louie, M. W. Khaw, and P. W. Glimcher, "Normalization is a general neural mechanism for context-dependent decision making," *Proc Natl Acad Sci U S A,* vol. 110, no. 15, pp. 6139-44, Apr 9 2013.

[150] L. Busse, A. R. Wade, and M. Carandini, "Representation of concurrent stimuli by population activity in visual cortex," (in eng), *Neuron,* vol. 64, no. 6, pp. 931-42, Dec 24 2009.

[151] F. J. Chaure, H. G. Rey, and R. Quian Quiroga, "A novel and fully automatic spike-sorting implementation with variable number of features," *J Neurophysiol,* vol. 120, no. 4, pp. 1859-1871, Oct 1 2018.

[152] A. Asadollahi, S. P. Mysore, and E. I. Knudsen, "Rules of competitive stimulus selection in a cholinergic isthmic nucleus of the owl midbrain," (in eng), *J Neurosci,* vol. 31, no. 16, pp. 6088-97, Apr 20 2011.

[153] M. R. Roesch and C. R. Olson, "Neuronal activity in primate orbitofrontal cortex reflects the value of time," *J Neurophysiol,* vol. 94, no. 4, pp. 2457-71, Oct 2005.

[154] J. W. Kable and P. W. Glimcher, "The neural correlates of subjective value during intertemporal choice," (in eng), *Nat Neurosci,* vol. 10, no. 12, pp. 1625-33, Dec 2007.

[155] C. Padoa-Schioppa and J. A. Assad, "Neurons in the orbitofrontal cortex encode economic value," *Nature,* vol. 441, no. 7090, pp. 223-6, May 11 2006.

[156] C. K. Machens, R. Romo, and C. D. Brody, "Functional, But Not Anatomical, Separation of "What" and "When" in Prefrontal Cortex," (in English), *Journal of Neuroscience,* vol. 30, no. 1, pp. 350-360, Jan 6 2010.

[157] C. A. Seger and E. K. Miller, "Category learning in the brain," *Annu Rev Neurosci,* vol. 33, pp. 203-19, 2010.

[158] M. D. Humphries, R. D. Stewart, and K. N. Gurney, "A physiologically plausible model of action selection and oscillatory activity in the basal ganglia," *J Neurosci,* vol. 26, no. 50, pp. 12921-42, Dec 13 2006.

[159] L. Wang and R. J. Krauzlis, "Visual Selective Attention in Mice," *Curr Biol,* vol. 28, no. 5, pp. 676-685 e4, Mar 5 2018.

[160] G. Felsen and Z. F. Mainen, "Neural substrates of sensory-guided locomotor decisions in the rat superior colliculus," (in eng), *Neuron,* vol. 60, no. 1, pp. 137-48, Oct 9 2008.

[161] W. T. Newsome, K. H. Britten, and J. A. Movshon, "Neuronal correlates of a perceptual decision," *Nature,* vol. 341, no. 6237, pp. 52-4, Sep 7 1989.

[162] V. Mante, D. Sussillo, K. V. Shenoy, and W. T. Newsome, "Context-dependent computation by recurrent dynamics in prefrontal cortex," *Nature,* vol. 503, no. 7474, pp. 78-84, Nov 7 2013.

[163] W. Chaisangmongkon, S. K. Swaminathan, D. J. Freedman, and X. J. Wang, "Computing by Robust Transience: How the Fronto-Parietal Network Performs Sequential, Category-Based Decisions," *Neuron,* vol. 93, no. 6, pp. 1504-1517 e4, Mar 22 2017.

[164] J. P. Fadok *et al.*, "A competitive inhibitory circuit for selection of active and passive fear responses," *Nature,* vol. 542, no. 7639, pp. 96-100, Feb 2 2017.

[165] T. Hafting, M. Fyhn, S. Molden, M. B. Moser, and E. I. Moser, "Microstructure of a spatial map in the entorhinal cortex," *Nature,* vol. 436, no. 7052, pp. 801-6, Aug 11 2005.

[166] B. W. Towse, C. Barry, D. Bush, and N. Burgess, "Optimal configurations of spatial scale for grid cell firing under noise and uncertainty," *Philos Trans R Soc Lond B Biol Sci,* vol. 369, no. 1635, p. 20130290, Feb 05 2014.

[167] D. Mao, S. Kandler, B. L. McNaughton, and V. Bonin, "Sparse orthogonal population representation of spatial context in the retrosplenial cortex," *Nat Commun,* vol. 8, no. 1, p. 243, Aug 15 2017.

[168] B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: a strategy employed by V1?," *Vision Res,* vol. 37, no. 23, pp. 3311-25, Dec 1997.

[169] J. B. Kruskal, Jr., "On the shortest spanning subtree of a graph and the traveling salesman problem," *Proc Amer Math Soc,* vol. 7, pp. 48-50, 1956.

[170] E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan, and D. B. Shmoys, "The traveling salesman problem. A guided tour of combinatorial optimization," *J Oper Res Soc,* vol. 37, pp. 536-536, 1985.

[171] Z. D. Jiang, A. J. King, and D. R. Moore, "Topographic organization of projection from the parabigeminal nucleus to the superior colliculus in the ferret revealed with fluorescent latex microspheres," (in eng), *Brain Res,* vol. 743, no. 1-2, pp. 217-32, Dec 16 1996.

[172] A. M. Treisman and G. Gelade, "A feature-integration theory of attention," *Cogn Psychol,* vol. 12, no. 1, pp. 97-136, Jan 1980.

[173] M. Usher and J. L. McClelland, "The time course of perceptual choice: the leaky, competing accumulator model," (in eng), *Psychol Rev,* vol. 108, no. 3, pp. 550-92, Jul 2001.

[174] A. K. Churchland and J. Ditterich, "New advances in understanding decisions among multiple alternatives," (in English), *Current Opinion in Neurobiology,* vol. 22, no. 6, pp. 920-926, Dec 2012.

[175] R. M. Roe, J. R. Busemeyer, and J. T. Townsend, "Multialternative decision field theory: A dynamic connectionist model of decision making," (in English), *Psychological Review,* vol. 108, no. 2, pp. 370-392, Apr 2001.

[176] R. Ratcliff and G. McKoon, "The diffusion decision model: theory and data for two-choice decision tasks," (in eng), *Neural Comput,* vol. 20, no. 4, pp. 873-922, Apr 2008.

[177] K. F. Wong and X. J. Wang, "A recurrent network mechanism of time integration in perceptual decisions," *J Neurosci,* vol. 26, no. 4, pp. 1314-28, Jan 25 2006.

[178] H. M. Schryver and S. P. Mysore, "Spatial Dependence of Stimulus Competition in the Avian Nucleus Isthmi Pars Magnocellularis," *Brain Behav Evol,* vol. 93, no. 2-3, pp. 137-151, 2019.

[179] H. M. Schryver, M. Straka, and S. P. Mysore, "Categorical signaling of the strongest stimulus by an inhibitory midbrain nucleus," *bioRxiv,* 2019.

[180] K. Zhang and T. J. Sejnowski, "Neuronal tuning: To sharpen or broaden?," *Neural Comput,* vol. 11, no. 1, pp. 75-84, Jan 1 1999.

[181] W. K. You and S. P. Mysore, "Endogenous and exogenous control of visuospatial selective attention in freely behaving mice," *bioRxiv,* 2019

[182] D. Silver *et al.*, "Mastering the game of Go without human knowledge," (in English), *Nature,* vol. 550, no. 7676, pp. 354-+, Oct 19 2017.

[183] A. Nguyen, J. Yosinski, and J. Clune, "Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images," (in English), *2015 Ieee Conference on Computer Vision and Pattern Recognition (Cvpr),* pp. 427-436, 2015.

[184] S. G. Finlayson, J. D. Bowers, J. Ito, J. L. Zittrain, A. L. Beam, and I. S. Kohane, "Adversarial attacks on medical machine learning," *Science,* vol. 363, no. 6433, pp. 1287-1289, Mar 22 2019.

[185] N. Mahajan, N. Mesgarani, and H. Hermansky, "Principal Components of Auditory Spectro-Temporal Receptive Fields," *Interspeech* p. 5, 2014.

[186] N. R. Mahajan, N. Mesgarani, and H. Hermansky, "General properties of auditory spectro-temporal receptive fields," *J Acoust Soc Am,* vol. 146, no. 6, p. EL459, Dec 2019.

[187] A. Mueller, D. S. Hong, S. Shepard, and T. Moore, "Linking ADHD to the Neural Circuitry of Attention," *Trends Cogn Sci,* vol. 21, no. 6, pp. 474-488, Jun 2017.

Curriculum Vitae

Nagaraj R. Mahajan

EDUCATION

- **PhD,** Electrical and Computer Engineering, Johns Hopkins University *Feb, 2020 (expected)*
- **MSE,** Electrical and Computer Engineering, Johns Hopkins University *May 2017*
- **Student,** Neural Systems and Behavior course, MBL Woods Hole *Aug 2016*
- **B. Tech (Hons.),** Electrical Engineering, IIT Hyderabad, India *May 2013*

RESEARCH WORK

- **Graduate research assistant,** Johns Hopkins University

Adviser: Dr. Shreesh P. Mysore *Apr 2014 - Mar 2020*
Thesis: Neural circuit mechanisms of target selection & attention

Adviser: Dr. Hynek Hermansky *Sep 2013 - Mar 2014*
Topic: Dominant encoding properties of ferret auditory cortex

- **Undergraduate research assistant,** IIT Hyderabad

Adviser: Dr. K Sri Rama Murty *Jan 2012 - May 2013*
Thesis**:** Speaker recognition system

Adviser: Dr. Sumohana Channappayya *Mar 2012 - May 2013*
Topic: Blood vessel count & area measurement in choroid scans

Adviser: Dr. Soumya Jana *August - December 2012*
Topic: Symmetry conditions in a Berger Tung problem.

- **Research intern, IISc Bangalore** *May-Jul 2012*

Adviser: Dr. Chandra Sekhar Seelamantula
Topic: Design of interpolated lookup tables

PUBLICATIONS

- **Mahajan NR**, Mysore SP. Neural circuit mechanism for generating categorical representations. *Under review.*
- Vandaele Y, **Mahajan NR**, Ottenheimer D, Richard JM, Mysore SP, Janak PH. Distinct recruitment of dorsomedial and dorsolateral striatal neural activity in an instrumental task erodes with extended training. *eLife, 2019.*
- **Mahajan NR**, Mesgarani N, Hermansky H. General properties of auditory spectro-temporal receptive fields. *JASA, 2019*
- **Mahajan NR**, Mysore SP. Combinatorial Neural Inhibition for Stimulus Selection across Space. *Cell Reports*, 2018
- **Mahajan NR**, Mesgarani N, Hermansky H. Principal components of auditory spectro-temporal receptive fields. *IEEE Interspeech*, 2014
- **Mahajan NR**, Donapati RCR, Channappayya SS, Vanjari S, Richhariya A, Chhablani J. An automated Algorithm for Blood Vessel Count and Area Measurement in 2-D Choroidal Scan Images. *IEEE EMBC*, 2013.

CONFERENCE PRESENTATIONS

- Mahajan NR, Mysore SP. Neural circuit strategy for categorical selection in the midbrain. *Cosyne, 2020*
- Mahajan NR, Mysore SP. Neural mechanisms of categorization for stimulus selection. *Society for Neuroscience*, 2017
- Mahajan NR, Mysore SP. Combinatorial inhibition for stimulus selection. *Johns Hopkins Bayview Research Symposium*, 2016
- Mahajan NR, Mysore SP. Combinatorial inhibition for efficient location-invariant stimulus selection. *Society for Neuroscience*, 2016
- Mahajan NR, Mysore SP. Atypical encoding of sensory space by an inhibitory nucleus in the midbrain attention network. *Society for Neuroscience*, 2015

TEACHING ASSISTANT

- *Fall 2015*: Gamechangers, Conceptual breakthroughs in neuroscience
- *Spring 2015, 2018*: Quantitative Methods for the Brain Sciences

AWARDS & DISTINCTIONS

- Collaborative research award 2016, Psychological & Brain sciences department, JHU.
- Stanley Watson scholarship 2016, to attend the NS & B course at MBL.
- ECE Graduate Fellowship 2013, Johns Hopkins University.
- Student travel grant, IEEE Interspeech 2014, Singapore.
- JENESYS scholarship, 2010 for an industrial visit sponsored by the Japanese government, for excellence in academics.
- All India ranks 2800 and 2205 (Top 1 %) in IIT-JEE and AIEEE 2009 examinations respectively.

POSITIONS HELD

- Volunteer for the JHU chapter of Association for India's development (AID), a non-profit organization that supports grassroots organizations in India.
- Cofounded and served as the first coordinator of the Electronics Club at IIT Hyderabad. Presented lectures, organized hands-on hobby electronics projects for the student body of IIT Hyderabad.
- Infrastructure coordinator at Elan 2011 (Annual techno-cultural fest of IIT Hyderabad).
- Served as a member of the Science and technology council of IIT-Hyderabad.
- Member of the IIT Hyderabad Basketball and Badminton teams.